面向鲁棒激光雷达语义分割的自适应增强感知潜在学习

Wangkai Li[1], **Zhaoyang Li**[1], **Yuwen Pan**[1],**Rui Sun**[1],**Yujia Chen**[1],**Tianzhu Zhang**[1,2*]中国科学技术大学[2]深空探测国家重点实验室、深空探测实验室{lwklwk, lizhaoyang, panyw, issunrui, yujia chen}@mail.ustc.edu.cn, tzzhang@ustc.edu.cn

arXiv:2603.01074v1 [cs.CV] 1 Mar 2026

摘要

恶劣天气条件会造成大幅分布偏移，显著降低激光雷达点云语义分割网络的性能。现有的基于增强的方法尝试在训练期间模拟天气干扰，以提升鲁棒性。然而，由于轻度增强和激进增强之间存在权衡，这些方法难以充分发挥数据增强的潜力。为此，我们提出 A3Point，这是一种自适应增强感知潜在学习框架，能够有效利用多种增强方式，同时减轻语义偏移，即增强导致语义含义发生变化的现象。A3Point 包含两个关键组成部分：语义混淆先验（SCP）潜在学习，用于捕获模型固有的语义混淆信息；以及语义偏移区域（SSR）定位，用于将语义混淆与语义偏移解耦，从而针对不同干扰程度采用自适应优化策略。在恶劣天气条件下，针对多个标准广义激光雷达分割基准进行的大量实验验证了我们方法的有效性，并取得了新的最先进结果。

1 引言

激光雷达语义分割对于自动驾驶等三维视觉任务至关重要（Li & Ibanez- Guzman, 2020；Li 等人, 2020；Aksoy 等人, 2020；Zhao 等人, 2023）。然而，现有方法（Ando 等人, 2023；Choy 等人, 2019；Lai 等人, 2023；Puy 等人, 2023）在恶劣天气（如雾、雪和雨）下往往表现不佳，这是因为点云发生严重的分布偏移，导致训练数据与测试数据不匹配。由于大多数户外场景点云数据集（Behley 等人, 2019；Fong 等人, 2022；Xiao 等人, 2022b）是在正常天气下采集的，因此，开发能够适应各种环境的鲁棒网络变得愈发重要。应对这一挑战对于实现可靠且天气不变的激光雷达语义分割至关重要。

为缓解性能下降，现有研究（Xiao 等，2023；Park 等，2024；Kong 等，2023b）通过基于仿真或基于增强的方法，在训练期间模拟恶劣天气，从而提升网络的鲁棒性。基于仿真的方法（Bijelic 等人，2018；Hahner 等人，2022；2021）通过建立物理方程模型来复现天气对点云的影响，但每种天气条件都需要单独建模，因此难以实际覆盖所有变化。基于增强的方法（Xiao 等，2023；Park 等，2024；Kim 等，2023）通过引入几何扰动和点丢弃，更灵活地模拟天气引起的失真。然而，这些方法的应用仍不充分，原因有二（图 1）：（1）轻度增强难以推广至严重天气条件；（2）过度增强会扭曲点云分布，导致语义偏移（Wang 等，2021；Bai 等，2022；Yuan 等，2021），使增强区域不再与原始语义一致，从而阻碍训练。这一困境促使现有方法限制增强范围和幅度，制约了其潜力。如何在缓解语义偏移的同时利用更大的增强空间，仍是一项引人关注的挑战。

*通讯作者

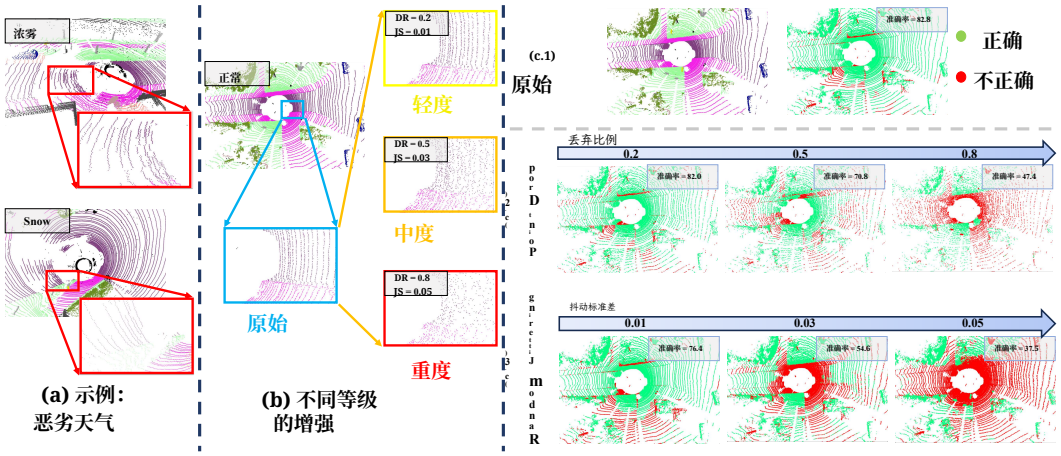
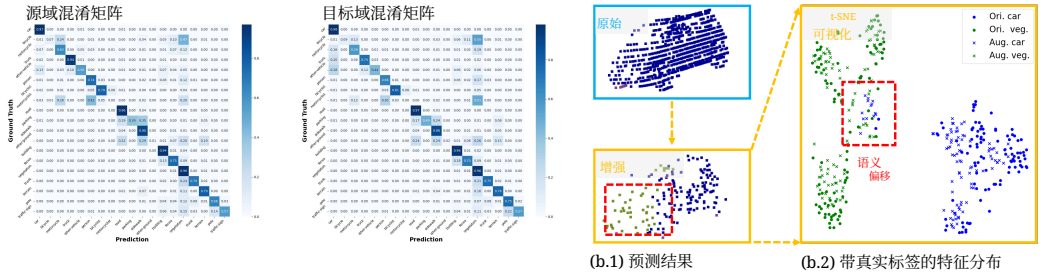


图 1：（a）恶劣天气条件造成的点云失真。（b）不同程度的增强（轻度、中度和重度），其中我们调整点丢弃的丢弃比例（DR）以及随机抖动的抖动标准差（JS）。（c）不同增强程度下分割准确率的可视化结果，表明剧烈失真会导致显著的性能下降。

这一分析促使我们探索范围更广、力度更大的增强空间，以更好地模拟不同强度下天气引起的失真。然而，要确保其有效性，就必须解决增强点云中可能出现的语义偏移问题。直接对语义偏移进行建模并不容易，因为预测错误源于两个因素：（1）**语义混淆**（图 2（a）），这是网络固有的特性，即使标签正确，网络也难以区分相似类别（例如，道路 与人行道）；（2）**语义偏移**（图 2（b）），即过度增强扭曲了点云分布，导致原始标签无法描述相应区域。对于语义混淆，应保留原始标签，以增强网络的判别能力。相反，对于语义偏移，则需要调整监督信号，避免产生误导。因此，解决语义偏移问题的关键在于将增强点云中的这两个因素区分开来。

本文对增强点云中的语义偏移进行建模，以便在训练过程中有效利用它们。我们观察到，原始数据和增强数据中均存在语义混淆，且这种混淆在不同域之间保持一致；而语义偏移仅出现在增强数据中。基于此，我们提出一种两步策略来识别语义偏移：（1）从正常条件下的预测中挖掘语义混淆先验。受 VQVAE（Van Den Oord 等人，2017）的启发，我们将其构建为**离散潜在表示学习任务**：将类别特定的局部混淆模式编码到潜在空间，并以量化潜变量表示，同时通过重建来约束表示能力。（2）将增强点云中的语义偏移检测为异常。我们将学习到的潜在编码器应用于增强预测，并将语义偏移定位构建为**异常检测问题**：通过将增强潜在表示与学习到的先验进行比较，我们区分语义一致性区域（仅受语义混淆影响）与语义偏移区域（还受到语义偏移影响）。这种区分使训练期间的定向优化成为可能。

基于上述讨论，我们提出面向恶劣条件下鲁棒激光雷达语义分割的**A**daptive**A**ugmentation-**A**ware潜在学习，即A3Point，其包含两个关键组件：语义混淆先验（SCP）潜变量学习与语义偏移区域（SSR）定位，以充分探索大规模且多样化增强空间在鲁棒点云分割中的潜力。**为捕捉网络固有的语义混淆**，在SCP潜在学习模块中，我们对来自原始点云的预测进行逐类别潜在编码，并使用量化潜变量表示局部混淆模式。通过向量量化与重建约束，该过程学习到能够重建预测图的有意义且具代表性的嵌入。**为将语义混淆与语义偏移解耦**，在SSR定位模块中，我们动态跟踪每个量化潜变量的表征分布，并使用冻结的先验潜在编码器隐式表示增强预测。通过将语义偏移检测视为异常检测问题，我们自适应地区分语义一致性区域（SCR）与语义偏移区域（SSR）。在SCR中，原始标签对优化仍然有效；而在SSR中，我们应用知识蒸馏，选择全局最近量化潜变量作为监督约束。通过联合定位SSR并自适-



(a) 语义混淆　　(b) 语义偏移

图 2：语义混淆与语义偏移示例。(a) 源域（正常天气）和目标域（恶劣天气）的混淆矩阵。尽管存在域偏移，语义混淆仍保持一致。(b) 激进增强会改变点云密度和形状，导致语义错位（例如，汽车 $\rightarrow$ 植被。）。

应地相应调整优化策略，A3Point充分借助多样化的数据增强，以提升不同扰动水平下的分割鲁棒性。

我们的贡献如下：1）我们提出一种全新的视角，以克服基于增强的方法的局限性，从而有效利用大规模且多样化的增强空间，提升激光雷达分割在恶劣天气下的鲁棒性。2）我们提出一个两步框架，将语义混淆解耦并减轻语义偏移，该框架包括语义混淆先验（SCP）潜在学习和语义偏移区域（SSR）定位。3）我们在多个标准泛化激光雷达分割基准上，针对恶劣天气开展了广泛实验，验证了我们的方法，并取得了新的最先进的结果。

2 相关工作

2.1 三维点云语义分割

三维点云的语义分割为每个点分配一个标签，通常可通过三种范式实现：基于点的方法、基于投影的方法和基于体素的方法。基于点的方法直接将三维点作为输入。PointNet（Qi 等人，2017）首次引入了这一方法，利用多层感知机提取逐点特征。后续研究（Thomas 等人，2019；Zhao 等人，2021；Choe 等人，2022；Fan 等人，2021；Li 等人，2024；2025e）进一步推动了这一范式的发展。尽管这些方法能够最大限度地减少信息损失并取得较强的性能，但将其应用于大规模激光雷达数据时，需要消耗大量计算资源。基于投影的方法（Ando 等人，2023；Kong 等人，2023a；Milioto 等人，2019；Xiao 等人，2021；Zhang 等人，2020）通过球面投影将点云映射为二维距离视图图像，从而能够使用二维分割网络。这些方法具有较高的计算效率，但在投影过程中会产生信息损失，导致准确率略低。基于体素的方法（Choy 等人，2019；Lai 等人，2023；Zhu 等人，2021；Graham 等人，2018；Tang 等人，2020）将三维点云划分为稀疏体素网格，并聚合同一体素内的点。稀疏卷积的引入显著降低了计算成本，使该范式非常适用于大规模户外激光雷达场景。在本文中，我们采用基于体素的架构（Choy 等人，2019；Tang 等人，2020）作为基线，以平衡推理效率和分割性能。

2.2 恶劣天气条件下的激光雷达

在实际应用中，恶劣天气下的鲁棒场景理解至关重要，尤其是考虑到自动驾驶的安全要求（Kong 等人，2023b；Xiao 等人，2023；Sakaridis 等人，2021；Pan 等人，2025）。然而，极端天气会显著扰乱点云分布（Filgueira 等人，2017；Heinzler 等人，2019；Peynot 等人，2009；Ryde 和 Hillier，2009），导致激光雷达分割网络的性能大幅下降。为缓解正常天气条件与恶劣天气条件之间的域差异，人们探索了无监督域自适应（UDA）方法（Hahner 等人，2022；2021；Luo 等人，2021；Xiao 等人，2022a；b；Yang 等人，2021）。受半监督学习的启发（Sun 等人，

2025年a；c；b），这些方法利用带标签的源域数据和未标注的目标域数据来学习域不变特征，从而提升跨域性能（Li 等人，2025年b；c；a；d；2026；Chen 等人，2025年b；a；He 等人，2025年a；b）。然而，基于无监督域自适应的方法仅适用于特定且可见的目标域，因此不足以泛化到未知的天气干扰。本文采用域泛化设置（Kim 等人，2024；2023；Li 等人，2023a；b），旨在仅使用正常天气条件下的单源域训练模型，且训练期间无法获取恶劣天气下的目标域数据。

2.3 鲁棒激光雷达分割的增强

为提升激光雷达分割模型的鲁棒性，现有方法在训练过程中引入点云损坏，以模拟恶劣天气造成的干扰。基于模拟的方法（Bijelic 等人，2018；Hahner 等人，2022；2021）依赖先验天气知识构建物理模型，以人工模拟恶劣条件下的点云。不同于显式建模特定天气影响，基于增强的方法（Xiao 等人，2023；Park 等人，2024；Kim 等人，2023）提供了一种更通用、灵活的方案。受二维图像增强启发，基于混合的方法（Kong 等人，2023c；Nekrasov 等人，2021；Xiao 等人，2022a；Zhao 等人，2024）将两次激光雷达扫描进行混合，以提高训练多样性。为更好地模拟恶劣天气造成的扰动，近期研究（Park 等人，2024；Xiao 等人，2023）指出了两种主要的退化模式：（1）几何扰动；（2）点丢弃。PointDR（Xiao 等人，2023）主要提出了 SemanticSTF 基准以及与之配套的基于增强的基线方法，该方法通过几何扰动和点丢弃模拟天气引起的扰动。LiDARWeather（Park 等人，2024）提出了一种可学习的点丢弃策略，用于自适应增强。然而，这些方法仍受限于有限的扰动空间。相比之下，我们的方法探索了更广泛的扰动范围，并明确处理激进增强造成的语义偏移问题，从而实现更鲁棒的网络训练。

3 方法

3.1 问题定义

在用于激光雷达语义分割的域泛化（DG）中，网络在带标签的源域数据上进行训练，并需要泛化到未见过的目标域数据。具体而言，源域可表示为 $D_s = \{(x_i^S,\ y_i^S)\}_{i=1}^{N_S}$，其中 $x_i^S \in X_S$ 表示一次激光雷达点云扫描，$y_i^S \in Y_S$ 为其对应的逐点独热标签，涵盖 C 个类别。目标域可表示为 $D_t = \{(x_i^T)\}_{i=1}^{N_T}$，其中目标标签 Y_T 与 Y_S共享相同的标签空间。由于训练过程中无法访问目标域，为简洁起见，以下记号中省略上标 S/T。

3.2 预备知识

基于增强的域泛化训练。现有方法探索天气引起的扰动，并将其用作数据增强。这种方法可视为一种域随机化范式，用于学习域泛化网络。在训练过程中，首先在原始扫描上计算损失，以训练神经网络 f:

$$\mathcal{L}_{ce} = \frac{1}{N_S}\sum_{i=1}^{N_S}\frac{1}{n_i}\sum_{j=1}^{n_i}\ell_{ce}(f(x_{ij}), y_{ij}), \tag{1}$$

其中，ℓ_{ce} 表示逐体素交叉熵损失，n_i 表示 x_i中的有效体素数量。然后，对于每个 x_i，同时对（x_i, y_i）应用随机增强，得到（$\hat{}x_i, \hat{}y_i$）$=\mathcal{A}\{$（x_i, y_i）$\}$。增强训练对也通过交叉熵损失实现:

$$\hat{\mathcal{L}}_{ce} = \frac{1}{N_S}\sum_{i=1}^{N_S}\frac{1}{\hat{n}_i}\sum_{j=1}^{\hat{n}i}\ell_{ce}(f(\hat{x}_{ij}), \hat{y}_{ij}). \tag{2}$$

$\hat{}$ 总训练损失可表示为：$\mathcal{L} = \mathcal{L}_{ce} + \hat{\mathcal{L}}_{ce}$。

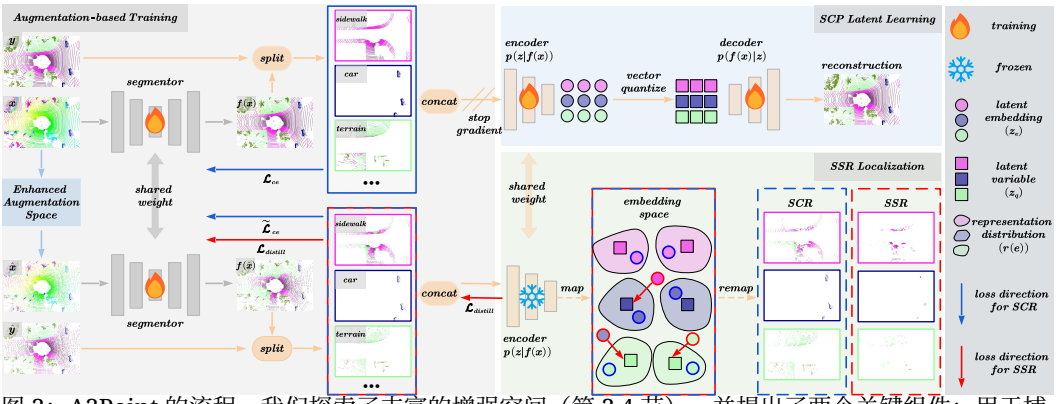


图 3：A3Point 的流程。我们探索了丰富的增强空间（第 3.4 节），并提出了两个关键组件：用于捕捉内在语义混淆的 SCP 潜变量学习（第 3.5 节），以及用于解耦语义偏移的 SSR 定位（第 3.6 节）。

向量量化变分自编码器。VQ-VAE（Van Den Oord 等人，2017）是变分自编码器（Kingma, 2013）的一种变体，可学习离散潜在表示。它由编码器 $\mathbb{E}$、解码器 $\mathbb{D}$ 和码本 $\mathcal{C} = \{e_1, e_2, ..., e_K\}$ 组成，码本中包含 K 个可学习的嵌入。编码器将输入 x 映射为连续潜在表示 $z_e = \mathbb{E}(x)$，随后使用最近邻查找将其量化为最近的码本条目 e_k（为简化起见，我们使用单个随机变量 z 表示离散潜变量）：

$$z_q = \text{quantize}(z_e) = e_k, \text{where } k = \arg\min_j ||z_e - e_j||_2. \tag{3}$$

解码器根据量化潜在表示重建输入：$\bar{x} = \mathbb{D}(z_q)$。总训练目标为：

$$\mathcal{L} = ||x - \bar{x}||_2^2 + ||\text{sg}(z_e) - z_q||_2^2 + ||z_e - \text{sg}(z_q)||_2^2, \tag{4}$$

其中包括重建损失、码本损失和承诺损失，sg(·) 表示停止梯度操作。码本损失使选定的潜变量 e 接近编码器输出，而承诺损失则促使编码器生成接近码本条目的潜在表示。
VQ-VAE 学习到的离散潜在表示能够捕捉数据的底层结构，并已成功应用于多种任务，例如图像生成（Razavi et al., 2019；De Fauw et al., 2019；Yu et al., 2021）和无监督表征学习（Liu et al., 2023；Chen et al., 2024；Takida et al., 2023）。

3.3 A3Point框架概述

我们首先定义训练期间使用的增强型增强空间（第3.4节）。然后，我们对离散潜在表示学习过程进行建模，以学习语义混淆先验（第3.5节）。接着，我们通过一种异常检测方法定位语义偏移区域（第3.6节）。最后，我们介绍区域特定优化策略（第3.7节）。整体框架如图3所示。

3.4 增强型增强空间

先前的研究（Xiao等，2023；Park 等人，2024）指出，恶劣天气造成的主要干扰可归纳为：（1）由感知距离偏移导致的几何扰动；（2）由光束衰减、光束缺失或潜在遮挡造成的点丢弃。这促使我们采用随机抖动和点丢弃作为主要的通用增强策略。不同于以往仅采用有限范围和幅度的方法，我们定义了更广泛的增强空间，以充分模拟不同程度的天气干扰。具体而言，我们将随机抖动的抖动标准差范围设为 $[j_{min}, j_{max}]$，将点丢弃的丢弃比例范围设为 $[d_{min}, d_{max}]$，并在训练期间对扰动幅度进行均匀采样。对于其他辅助增强，我们沿用先前的研究，采用随机旋转、随机缩放、随机翻转、随机噪声扰动和扫描混合（Kong 等人，2023c；Xiao 等，2022a）。

3.5 语义混淆先验潜在学习

语义混淆是指网络在区分类别时固有的不确定性，这种不确定性体现在预测概率分布中。为了将增强点云中的语义混淆与语义偏移解耦，我们从原始点云预测中挖掘语义混淆的先验知识。

具体而言，我们首先引入一种用于先验潜变量学习的自动编码器，该编码器遵循 VQ-VAE（Van Den Oord 等人，2017）。通过量化潜变量和重建过程，它能够有效学习具有足够表征能力的有意义表征。编码器$\mathbb{E}$ 的输入通过以下方式获得：先将 f（x经过 softmax 处理）与 x拼接，再根据标签 y 按类别拆分结果，以防止类别间交互，最后沿批次维度拼接所得子矩阵。可表示为：

$$z_e = \mathbb{E}(\underbrace{[f(x^1) \oplus x^1, f(x^2) \oplus x^2, ..., f(x^C) \oplus x^C]}_{n\times(C+3)}), \tag{5}$$

其中，f（x^i）$\oplus x^i$ 表示类别为i的点的预测结果与坐标的拼接，n是有效体素的数量。符号 $[\cdot,\cdot]$表示沿批次维度进行拼接的操作。解码器 $\mathbb{D}$的输出被重建为 $[f$（x^1）；f（x^2）；...；f（x^C）$]$。使用公式 4 对这一过程进行优化，且不将梯度反向传播至 f。

与 VQ-VAE 中的 $K \times D$ 代码码本相比，我们使用大小为 $C \times k \times D$ 的类别专属子码本，其中 k 是每个子码本中的潜变量数量，D 是潜变量的维度。子码本中的每个潜变量都可以视为对相应类别下特定局部语义分布模式的建模。编码器 $\mathbb{E}$ 对 p（$z|f$（x））进行建模，表示给定预测概率 f（x）时潜变量 z 的先验分布。解码器 $\mathbb{D}$ 对 p（f（x）$|z$）进行建模，表示以潜变量 z 为条件重建预测概率 f（x）的后验分布。通过这一过程，自动编码器可以在训练过程中在线建模网络的语义混淆。

3.6 语义偏移区域定位

过度增强可能导致某些区域发生语义偏移，从而产生异常预测结果，如图 2 (b) 所示。在对语义混淆先验建模后，我们可以评估网络对增强后的点云所做的预测是否符合正常语义分布模式，并将语义偏移区域的定位视为异常检测问题。

在潜变量学习过程中，我们动态跟踪每个量化潜变量的表征分布。对于每个 e_i，我们维护其量化前对应潜在嵌入分布的统计量，并存储每个通道的方差 σ_i^2，该方差使用指数移动平均进行更新。每个 e_i 的表征分布可表示为：

$$r(e_i) = \mathcal{N}(e_i, \mathrm{diag}(\sigma_{i,1}^2, \sigma_{i,2}^2, \ldots, \sigma_{i,D}^2)), \tag{6}$$

接下来，我们使用冻结的先验潜在编码器 $\mathbb{E}$ 将增强点云的预测映射到潜嵌入空间。由于该编码器对类别级 p（$z|f$（x））进行建模，受语义偏移影响区域的嵌入不会落入其对应子码本的表征分布中。我们通过检查这些嵌入是否位于对应子码本中其最近潜变量的表征分布内，来定位语义偏移区域。

在区分所有嵌入之后，我们将它们重新映射回预测空间，即 f(ˆx)，以确定预测空间中的语义一致性区域（SCR）和语义偏移区域（SSR）。我们使用两个掩码来表示这两个区域：

$$M_{SCR} = \mathbb{1}(z_e \in r(NN_{sub}(z_e))), \tag{7}$$

$$M_{SSR} = \mathbb{1}(z_e \notin r(NN_{sub}(z_e))), \tag{8}$$

其中，z_e 表示增强点云预测的潜在嵌入，NN_{sub}(z_e)表示 z_e的最近潜变量，而$\mathbb{1}$(·)是指示函数。

3.7 优化策略

确定 SCR 和 SSR 后，我们采用不同的优化策略。对于 SCR，我们使用原始标签：

$$\widetilde{\mathcal{L}}_{ce} = \ell_{ce}(f(\hat{x}) \odot M_{SCR}, y \odot M_{SCR}) = \hat{\mathcal{L}}_{ce} \odot M_{SCR}, \tag{9}$$

其中 $\odot$ 表示逐元素乘法。对于 SSR，我们提出一种基于潜变量的蒸馏损失，以提供适当的监督信号。我们不从对应类别的子码本中查询最近邻，而是从全局码本中查询，获取该区域最接近的语义混淆模式先验。然后，我们使用此先验来蒸馏 SSR 的潜在嵌入：

$$\mathcal{L}_{distill} = ||z_e - sg[NN_{global}(z_e)]||_2^2, \tag{10}$$

其中 $NN_{global}(z_e)$ 表示全局最近邻潜变量。公式 10 与以往的承诺损失类似，但我们不更新编码器 $\mathbb{E}$，而是将梯度反向传播至网络 f。我们的总训练损失为：

$$\mathcal{L}_{total} = \mathcal{L}_{ce} + \widetilde{\mathcal{L}}_{ce} + \lambda \mathcal{L}_{distill}, \tag{11}$$

其中 λ 是用于平衡各项损失的超参数。通过这种方式，从原始点云中学习到的语义混淆先验可以为语义偏移区域提供有意义的监督信号，从而提升模型在增强数据上的性能，同时保持语义一致性。有关更多实现细节、讨论和算法流程，请参阅附录 B-E。

4 实验

4.1 实验设置

数据集。 我们使用四个数据集：SemanticKITTI（Behley 等人，2019）、SynLiDAR（Xiao 等人，2022b）、SemanticKITTI-C（Kong 等人，2023b）和 SemanticSTF（Xiao 等，2023）。❶ SemanticKITTI：19,130 次训练扫描（序列 00–10，不包括 08），采集于标准天气条件下的城市环境中。❷ SynLiDAR：使用虚幻引擎 4 生成的 198,396 次合成扫描（190 亿个点）。❸ SemanticKITTI-C：通过模拟生成的 SemanticKITTI 损坏版本。❹ SemanticSTF：来自 STF（Bijelic 等人，2020）的 2,076 次激光雷达扫描，采集于恶劣天气（雪、浓雾、轻度雾、雨）下，划分为 1,326 次训练扫描、250 次验证扫描和 500 次测试扫描。SemanticKITTI 和 SynLiDAR 作为源域，而 SemanticKITTI-C 和 SemanticSTF 则作为目标域，用于评估鲁棒性。为简洁起见，我们将 SemanticKITTI 记为 [A], , 将 SynLiDAR 记为[B], , 将 SemanticSTF 记为 [C], , 将 SemanticKITTI-C 记为 [D] 。

评估指标。 我们采用 MinkowskiNet-18/32width（Choy 等人，2019）作为基础模型，并同时评估 SPVCNN（Tang 等人，2020）。性能以每个类别的交并比以及各类别的平均交并比来衡量，并按天气条件进行细分。

实现细节。 分割网络使用随机梯度下降进行训练（学习率 0.24，权重衰减 0.0001），自动编码器使用 Adam 进行训练（学习率 0.001）。子码本大小 k 设为 32。这些网络交替更新。对于增强，我们将丢弃比例设为 [0.2、0.8], 并将抖动标准差设为 [0.01、0.05]。训练在 RTX 3090（24 GB）上以批量大小 4 运行 50 个训练轮次。平衡系数 λ 设为 0.1，以保持梯度稳定性。

4.2 与现有方法的比较

总体定量结果。表 1-2 比较了 A3Point 与在 $[A] \rightarrow [C]$和 $[B] \rightarrow [C]$ 基准测试上表现领先的域泛化方法。基线仅使用 $\mathcal{L}_{ce}$进行训练。A3Point 优于所有方法，相较于基线，平均交并比分别提升了 9.9% 和 11.7%。按类别分析发现，A3Point 在摩托车、骑自行车者、交通标志和汽车等安全关键类别上表现尤为出色。这些类别通常因几何形状独特且容易受到恶劣条件影响而具有挑战性。VQ-VAE 中的局部模式编码和区域特定优化有助于在多种天气条件下为这些类别提取稳健特征。

天气等级对比。 如表 1-2 所示，A3Point 在所有天气条件下都展现出卓越的鲁棒性，即使在浓雾等最恶劣的情况下也大幅领先

表 1： [A] → [C] 的比较结果。 ∗ 表示使用相同骨干网络复现的结果。

方法	car	bi.cle	mt.cle	truck	oth-v.	pers.	bi.clst	mt.clst	road	parki.	sidew.	other-g.	build.	fence	veget.	trunk	terr.	pole	traf.	D-fog	L-fog	Rain	Snow	mIoU	增益
Oracle	89.4																							-	
基线	67.1	**5.0**	28.1	38.5	14.6	45.8	8.3	13.8	40.1	16.1	26.1	3.3	71.6	52.7	53.8	33.9	39.2	25.3	12.7	30.7	30.1	29.7	25.3	31.4	+0.0
PointDR* (Xiao等，2023)	69.2	1.0	8.9	**41.9**	7.6	48.9	17.0	36.2	57.8	15.9	32.3	4.0	75.7	46.4	54.0	36.2	43.9	23.7	24.2	37.3	33.5	35.5	26.9	33.9	+2.5
DGUIL（Kim 等，2023）	78.2	2.5	33.0	29.7	6.1	49.8	0.8	40.9	67.3	7.2	38.0	2.2	**79.8**	**54.4**	**64.1**	36.8	52.3	31.0	40.0	36.3	34.5	35.5	33.3	37.6	+6.2
WADG（Du 等，2024）	72.0	0.0	32.9	37.0	1.9	37.7	6.8	52.9	59.9	10.7	31.8	2.2	76.0	48.8	62.7	34.0	49.3	23.6	20.4	39.5	32.5	31.7	29.4	34.8	+3.4
DGLSS（He 等，2024）	69.6	0.8	42.8	34.4	8.9	41.9	12.8	44.5	52.0	14.5	30.8	**6.0**	77.8	51.1	57.6	38.9	43.2	29.7	30.6	34.2	34.8	36.2	32.1	36.2	+4.8
LiDARWeather（Park 等，2024）	86.1	4.8	13.8	39.7	**26.6**	**55.4**	8.5	50.4	63.7	14.9	37.9	5.5	75.2	52.7	60.4	39.7	44.9	30.1	40.8	36.0	37.5	37.6	33.1	39.5	+8.1
NTN（Park 等，2025）	83.3	3.7	31.3	36.2	18.2	53.3	6.8	**55.9**	67.2	**18.1**	37.2	5.4	72.1	41.8	58.0	36.0	46.0	28.2	39.8	35.3	35.1	35.7	32.4	38.9	+7.5
A3Point（我们的方法）	**88.3**																								

表2： [B] → [C]的比较结果。 ∗ 表示使用相同骨干网络复现的结果。

方法	car	bi.cle	mt.cle	truck	oth-v.	pers.	bi.clst	mt.clst	road	parki.	sidew.	other-g.	build.	fence	veget.	trunk	terr.	pole	traf.	D-fog	L-fog	Rain	Snow	mIoU	增益
Oracle	89.4																							-	
基线	33.8	1.7	3.3	15.5	0.2	25.5	1.6	3.4	15.3	9.2	16.8	0.1	33.4	21.9	39.5	18.7	**44.0**	8.8	0.8	15.2	16.0	16.8	12.8	15.5	+0.0
PointDR* (Xiao 等人，2023)	41.1	2.8	3.4	18.1	0.2	31.3	**2.8**	3.3	34.4	**10.2**	19.7	1.0	52.7	22.0	48.5	21.3	38.3	19.2	5.6	19.1	20.3	25.3	19.0	19.8	+4.3
WADG（Du 等，2024）	33.8	1.1	2.9	17.0	0.2	26.8	1.0	4.3	53.9	5.0	20.6	2.2	**64.3**	27.1	53.8	27.0	37.0	**28.6**	8.6	21.6	23.4	27.2	21.4	21.9	+6.4
LiDARWeather (Park等人,2024)	39.3	2.9	0.9	19.4	0.8	27.7	2.2	3.8	42.5	9.4	21.6	0.3	51.9	33.5	47.4	23.1	33.3	23.2	6.8	19.0	21.2	23.1	17.3	20.5	+5.0
NTN (Park等人, 2025)	48.4	1.5	2.4	19.4	0.2	29.1	3.2	8.9	43.5	6.7	20.5	0.0	52.2	30.1	49.8	20.0	32.9	24.7	7.5	-	-	-	-	21.1	+5.6
A3Point (本文)	**76.7**																								

以及大雪天气下也是如此。广泛的增强空间和自适应潜空间蒸馏使 A3Point 能够应对各种天气干扰，同时保留受损较轻区域的预测结果，因此在较温和的条件下也比其他方法表现更好。

定性结果。 图4展示了在雪和浓雾等具有挑战性的天气条件下，A3Point 与先前方法的可视化对比。A3Point 能够更准确地分割场景中的主要组成部分，例如人行道、道路和地形，而这些部分在恶劣天气下往往会被遮挡或扭曲。此外，它对汽车 和交通标志等复杂实例的分割效果也更好，而准确分割这些目标对于安全导航至关重要。局部模式编码有助于捕捉这些目标的独特几何特征，而对语义混淆与偏移进行解耦，则提升了这些类别的判别性能。

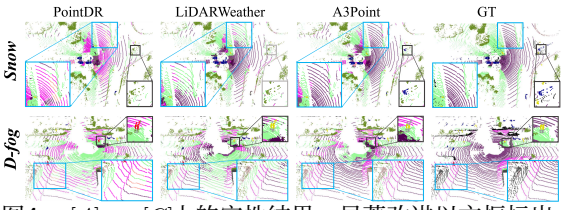


图4： $[A] \rightarrow [C]$上的定性结果。显著改进以方框标出。

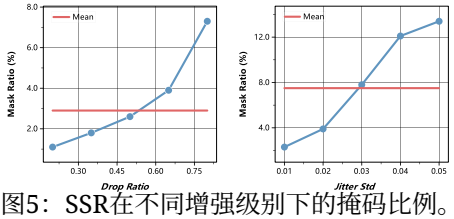


图5：SSR在不同增强级别下的掩码比例。

4.3 消融研究

更多分析、讨论和可视化结果请参见附录F-V。

更多结果。 为验证 A3Point 的有效性和泛化能力，我们在不同架构和基准测试上进行了实验（表3）。结果显示，性能均有提升。使用 SPVCNN 时，A3Point 在 $[A] \rightarrow [C]$、 $[B] \rightarrow [C]$ 和 $[A] \rightarrow [D]$ 上的平均交并比分别达到 +12.0%、 +8.5% 和 +2.2%。使用 Minkowski 时，平均交并比则分别提升了 +9.9%、 +11.7% 和 +5.9%。值得注意的是，在真实恶劣天气基准 $[C]$ 上的提升幅度大于在合成损坏基准 $[D]$ 上的提升幅度，这表明该方法对真实世界干扰具有更强的鲁棒性。不同架构上的稳定提升凸显了该方法对架构的友好性，使其能够广泛应用于各种激光雷达分割网络。

各组件的有效性。 我们在 [A]/[B]→[C] 基准测试上开展消融研究，以评估A3Point各组件的影响（表4）。对于 [A]→[C], , 不使用数据增强的基线取得了31.4%的平均交并比。引入增强型增强空间（EAS）后，性能提升至38.7%，凸显了多样化数据增强对域鲁棒性的重要性。应用语义偏移区域定位与掩码，仅结合 $\mathcal{L}_{ce}$ 优化语义一致性区˜域（SCR），将平均交并比进一步提高至40.2%。最后，引入基于潜变量的

表3：不同的架构与基准测试。

方法	SPVCNN		Minkowski	
	基线	使用 A3Point	基线	使用 A3Point
$[A] \rightarrow [C]$	28.1	40.1 (+12.0)	31.4	41.3 (+9.9)
$[B] \rightarrow [C]$	17.3	25.8 (+8.5)	15.5	27.2 (+11.7)
$[A] \rightarrow [D]$	52.5	54.7 (+2.2)	53.0	58.9 (+5.9)

桌子 4：A3Point 组件的消融研究。

无 EAS ($\hat{}\mathcal{L}_{ce}$)	SCR ($\tilde{\mathcal{L}}_{ce}$)	SSR ($\mathcal{L}_{distll}$)	$[A] \rightarrow [C]$	$[B] \rightarrow [C]$
✓			31.4	15.5
	✓		38.7	24.9
	✓	✓	40.2	26.5
	✓	✓	**41.3**	**27.2**

桌子 5：增强级别的消融研究。我们定义了不同级别（轻度、中度和重度），与图1相同。

方法	None	轻度	中度	重度	随机
$[A] \rightarrow [A]$	62.7	61.2	59.7	55.5	60.4
使用 A3Point	**62.6**	62.4	61.0	58.7	62.5
$[A] \rightarrow [C]$	31.4	37.6	38.0	37.1	38.7
搭配 A3Point	31.8	38.5	40.4	40.5	**41.3**
$[B] \rightarrow [C]$	15.5	19.3	23.6	24.1	24.9
搭配 A3Point	15.6	20.7	25.1	26.4	**27.2**

表6：不同语义混淆先验建模策略的性能比较。

策略	$[A] \rightarrow [C]$	$[B] \rightarrow [C]$
None	38.7	24.9
基于真实标注	39.1	25.0
离线	40.7	26.4
在线（我们的）	**41.3**	**27.2**

蒸馏损失 $\mathcal{L}_{distill}$ 用于语义偏移区域（SSR），取得了41.3%平均交并比的最佳结果。在 [B]→[C], 上观察到类似趋势，基线的平均交并比起始为15.5%。EAS将性能显著提升至24.9%，弥合了合成数据与真实恶劣天气数据之间的巨大差距。加入SCR后，平均交并比提升至26.5%，包含SSR的完整模型达到27.2%。这些结果表明，我们的各组件在域泛化场景中具有互补性。通过利用学习到的语义混淆先验来引导语义偏移优化，模型能够适应新的扰动，同时保持语义一致性。

增强级别分析。 我们使用 [A] 验证集（[A]→[A]）以及跨领域基准 [A]/[B]→[C] （表5）分析增强级别对网络性能的影响。在 [A]→[A], 上，随着增强从无增加到重度，基线性能的平均交并比从62.7%降至55.5%，表明严重数据增强具有不利影响。相比之下，A3Point保持了更高的性能，仅从62.6%略降至58.7%，表明其能够通过处理语义偏移区域来减轻激进增强的负面影响。对于 [A]→[C], ，数据增强将基线性能的平均交并比从31.4%提升至38.7%，证实了其在提升鲁棒性方面的作用。然而，从中度增强到强增强时，性能从38.0%略降至37.1%，表明开始出现语义偏移问题。A3Point避免了这一下降，在强增强下达到40.5%平均交并比，在随机增强下达到41.3%平均交并比。 [B]→[C], 中也出现了类似趋势：随着增强强度增加，A3Point取得了更大的相对增益。其相对于基线的性能差距从轻度增强时的1.4%扩大到随机增强时的2.3%。这些发现证实，A3Point能够有效处理语义偏移，从而可以在不损害性能的情况下使用激进增强。其平衡源域性能与目标域性能的能力使其非常适合实际应用。

语义混淆先验分析。 表6比较了在 [A]/[B]→[C] 基准上建模语义混淆先验的不同策略。对于 [A]→[C],, 基于真实标注的方法使用独热真实标签，仅建模类别级形状分布，但缺乏类间混淆知识，仅带来0.4%的mIoU增益。离线方法利用预训练模型在源域上的预测，将平均交并比提升了2.0%，但难以捕捉训练过程中不断演变的混淆模式。 [B]→[C],中也出现了类似趋势：基于真实标注的方法增益极小（25.0%对24.9%），离线方法虽有改善，但提升仍然有限（26.4%对24.9%）。相比之下，我们的在线建模方法会持续更新先验信息，从而有效解耦增强点云中的语义混淆。该动态策略取得了最佳性能：在 [A]→[C] 上平均交并比达到41.3%，在 [B]→[C],上达到27.2%，分别比基线高出2.6%和2.3%。通过持续适应不断演变的混淆模式，A3point缓解了语义偏移，并始终优于静态先验，证明了动态先验建模在我们框架中的重要性。

语义偏移区域分析。图5展示了 SSR 掩码比例如何随增强等级变化。随着增强强度增加，定位出的 SSR 掩码比例也随之增大。这一观察符合我们的预期：更强的扰动会导致增强点云产生更显著的语义偏移，从而使输入中被识别为属于 SSR 的部分占比更大。曲线呈上升趋势，说明我们提出的 SSR 定位模块能够有效捕捉受增强诱导的

语义偏移影响最大的区域，使网络能够在这些区域应用适当的优化策略，以减轻语义偏移的影响，并增强模型在恶劣天气条件下的鲁棒性。

5 结论

本文提出了 A3Point，这是一种用于恶劣天气下鲁棒激光雷达语义分割的自适应增强感知潜在学习框架。A3Point 通过两步策略有效利用庞大的增强空间，同时缓解语义偏移：（1）语义混淆先验潜在学习，通过离散潜在表示编码局部混淆模式；（2）语义偏移区域定位，在增强点云中检测异常，从而区分语义一致区域与语义偏移区域，实现定向优化。在域泛化基准上的实验表明，该方法效果显著，尤其是在从合成正常天气泛化到真实恶劣天气方面。

致谢

本研究部分得到了国家重点研发计划（项目编号：2024YFB3909902）、国家自然科学基金委员会（项目编号：U25A20536）以及中国科学院青年创新促进会的资助。

参考文献

Eren Erdal Aksoy、Saimir Baci 和 Selcuk Cavdar。Salsanet：面向自动驾驶的激光雷达点云快速道路与车辆分割。载于*2020*年*IEEE*智能车辆研讨会（*IV*），第 926–932 页。IEEE，2020。

Angelika Ando、Spyros Gidaris、Andrei Bursuc、Gilles Puy、Alexandre Boulch 和 Renaud Marlet。Rangevit：面向自动驾驶三维语义分割的视觉Transformer。载于*IEEE/CVF*计算机视觉与模式识别会议论文集，第 5240–5250 页，2023。

Yingbin Bai、Erkun Yang、Zhaoqing Wang、Yuxuan Du、Bo Han、Cheng Deng、Dadong Wang 和 Tongliang Liu。Rsa：通过激进增强减少自监督学习中的语义偏移。神经信息处理系统进展，35:21128–21141，2022。

Jens Behley、Martin Garbade、Andres Milioto、Jan Quenzel、Sven Behnke、Cyrill Stachniss 和 Jurgen Gall。《SemanticKITTI：用于激光雷达序列语义场景理解的数据集》。发表于*IEEE/CVF*国际计算机视觉会议论文集，第9297–9307页，2019年。

Mario Bijelic、Tobias Gruber 和 Werner Ritter。《雾中激光雷达传感器基准测试：检测是否正在失效？》。发表于*2018*年*IEEE*智能车辆研讨会 *(IV)*，第760–767页。IEEE，2018年。

Mario Bijelic、Tobias Gruber、Fahim Mannan、Florian Kraus、Werner Ritter、Klaus Dietmayer 和 Felix Heide。《不见雾，亦能穿雾：面向未见过的恶劣天气的深度多模态传感器融合》。发表于*IEEE/CVF*计算机视觉与模式识别会议，第11682–11692页，2020年。

Hao Chen、Ze Wang、Xiang Li、Ximeng Sun、Fangyi Chen、Jiang Liu、Jindong Wang、Bhiksha Raj、Zicheng Liu 和 EmadBarsoum。《Softvq-vae：高效的一维连续分词器》。*arXiv*预印本 *arXiv:2412.10958*，2024年。

Yujia Chen、Rui Sun、Wangkai Li、Huayu Mai、Si Chen、Zhuoyuan Li、Zhixin Cheng 和 Tianzhu Zhang。Beyondmix：利用结构先验和长程依赖实现域不变的激光雷达分割。发表于第三十九届神经信息处理系统年会，2025年a。

Yujia Chen、Rui Sun、Wangkai Li、Huayu Mai、Naisong Luo、Yuwen Pan 和 Tianzhu Zhang。缓解与挖掘：从伪标签视角重新思考线粒体分割的无监督域自适应。载于*AAAI*人工智能会议论文集，第 39 卷，第 2339–2347 页，2025年b。

Jaesung Choe、Chunghyun Park、Francois Rameau、Jaesik Park 和 In So Kweon。Pointmixer：用于点云理解的 MLP 混合器。载于欧洲计算机视觉会议，第 620–640 页。Springer，2022。

Christopher Choy、JunYoung Gwak 和 Silvio Savarese。4D 时空卷积网络：Minkowski 卷积神经网络。发表于*IEEE/CVF*计算机视觉与模式识别会议论文集，第 3075–3084 页，2019。

JeffreyDeFauw、Sander Dieleman 和 Karen Simonyan。带辅助解码器的分层自回归图像模型。*arXiv* 预印本 *arXiv:1903.04933*，2019。

Jing Du、John Zelek 和 Jonathan Li。天气感知自动驾驶系统：多种天气场景下点云语义分割的域泛化。*ISPRS* 摄影测量与遥感学报，218：204–219，2024。

Siqi Fan、Qiulei Dong、Fenghua Zhu、Yisheng Lv、Peijun Ye 和 Fei-Yue Wang。SCF-Net：面向大规模点云分割的空间上下文特征学习。发表于 *IEEE/CVF*计算机视觉与模式识别会议论文集，第 14504–14513 页，2021。

A Filgueira、H Gonz´alez-Jorge、Susana Lag¨uela、L D´ıaz-Vilari˜no 和 Pedro Arias。量化雨对激光雷达性能的影响。*Measurement*，95:143–148，2017。

Whye Kit Fong、Rohit Mohan、Juana Valeria Hurtado、Lubing Zhou、Holger Caesar、Oscar Beijbom 和 AbhinavValada。Panopticnuscenes：用于激光雷达全景分割和跟踪的大规模基准测试。*IEEE* 机器人与自动化快报，7(2):3795–3802，2022。

Benjamin Graham、Martin Engelcke 和 Laurens Van Der Maaten。使用子流形稀疏卷积网络进行三维语义分割。载于*IEEE*计算机视觉与模式识别会议论文集，第9224–9232页，2018年。

Martin Hahner、Christos Sakaridis、Dengxin Dai 和 Luc Van Gool。面向恶劣天气下三维目标检测的真实激光雷达点云雾模拟。载于*IEEE/CVF*国际计算机视觉会议论文集，第15283–15292页，2021年。

Martin Hahner、Christos Sakaridis、Mario Bijelic、Felix Heide、Fisher Yu、Dengxin Dai 和 Luc Van Gool。用于鲁棒三维目标检测的激光雷达降雪模拟。载于*IEEE/CVF*计算机视觉与模式识别会议论文集，第 16364–16374 页，2022 年。

Pei He、焦李成、Lingling Li、Xu Liu、Fang Liu、Wenping Ma、杨淑媛和尚荣华。面向域泛化的三维点云分割不确定性内省学习。载于第 *32* 届 *ACM*国际多媒体会议论文集，第 651–660 页，2024 年。

Weinan He、Zilei Wang 和 Yixin Zhang。利用文本表示进行目标语义聚类，实现鲁棒的通用域自适应。载于*AAAI*人工智能会议论文集，第 39 卷，第 17132–17140 页，2025年a。

Weinan He、Yixin Zhang 和 Zilei Wang。渐进分布桥接：通过自适应辅助数据为大规模预训练模型进行无监督适配。载于*IEEE/CVF*国际计算机视觉会议论文集，2025年b。

Robin Heinzler、Philipp Schindler、J¨urgen Seekircher、Werner Ritter 和 Wilhelm Stork。天气影响与车载激光雷达传感器分类。载于*2019 IEEE*智能车辆研讨会（*IV*），第1527–1534页。IEEE，2019。

Hyeonseong Kim、Yoonsu Kang、Changgyoon Oh 和 Kuk-Jin Yoon。激光雷达语义分割的单域泛化。载于*IEEE/CVF*计算机视觉与模式识别会议论文集，第17587–17598页，2023。

Jaeyeul Kim、Jungwan Woo、Jeonghoon Kim 和 Sunghoon Im。重新思考激光雷达域泛化：将单源视为多密度域。载于欧洲计算机视觉会议，第310–327页。Springer，2024。

Diederik P Kingma、Max Welling 等。自编码变分贝叶斯，2013。

Durk P Kingma 和 Prafulla Dhariwal。Glow：采用可逆1×1卷积的生成流。神经信息处理系统进展，31，2018。Lingdong Kong、Youquan Liu、Runnan Chen、Yuexin Ma、朱新革、Yikang Li、Yuenan Hou、Yu Qiao 和 Ziwei Liu。重新思考用于激光雷达分割的距离视图表示。载于*IEEE/CVF*国际计算机视觉会议论文集，第228–240页，2023a。Lingdong Kong、Youquan Liu、Xin Li、Runnan Chen、Wenwei Zhang、Jiawei Ren、Liang Pan、Kai Chen 和 Ziwei Liu。Robo3D：面向损坏情况下稳健可靠的三维感知。载于*IEEE/CVF*国际计算机视觉会议论文集，第19994–20006页，2023b。Lingdong Kong、Jiawei Ren、Liang Pan 和 Ziwei Liu。用于半监督激光雷达语义分割的Lasermix。载于*IEEE/CVF*计算机视觉与模式识别会议论文集，第21705–21715页，2023c。Xin Lai、Yukang Chen、Fanbin Lu、Jianhui Liu 和贾佳亚。用于基于激光雷达的三维识别的球面Transformer。载于*IEEE/CVF*计算机视觉与模式识别会议论文集，第17545–17555页，2023。Miaoyu Li、Yachao Zhang、Xu Ma、Yanyun Qu 和 Yun Fu。BEV-DG：面向三维语义分割域泛化的鸟瞰视图跨模态学习。载于*IEEE/CVF*国际计算机视觉会议论文集，第11632–11642页，2023a。Wangkai Li、Zhaoyang Li、Rui Sun、Huayu Mai、Naisong Luo、Yuan Wang、Yuwen Pan、Guoxin Xiong、Huakai Lai、Zhiwei Xiong 等。MAUNet：用于多模态细胞分割的模态感知抗歧义U-Net。载于神经信息处理系统竞赛，第1–12页。PMLR，2023b。Wangkai Li、Rui Sun、Zhaoyang Li 和 Tianzhu Zhang。迈向语义分割中稳健的伪标签学习：一种编码视角。载于第三十九届神经信息处理系统年会，2025年a。Wangkai Li、Rui Sun、Bohao Liao、Zhaoyang Li 和 Tianzhu Zhang。域自适应语义分割的平衡学习。载于国际机器学习会议，第35858–35883页。PMLR，2025年b。Wangkai Li、Rui Sun、Huayu Mai 和 Tianzhu Zhang。通过语义分割中的图像退化实现无监督域桥接。载于第三十九届神经信息处理系统年会，2025c。Wangkai Li、Rui Sun、Zhaoyang Li、Yujia Chen 和 Tianzhu Zhang。DA-CAL：迈向语义分割中的跨域校准。*arXiv*预印本 *arXiv:2602.20860*，2026。Ying Li、Lingfei Ma、Zilong Zhong、Fei Liu、Michael A Chapman、Dongpu Cao 和 Jonathan Li。自动驾驶中激光雷达点云的深度学习：综述。*IEEE*神经网络与学习系统汇刊，32(8):3412–3432，2020。You Li 和 Javier Ibanez-Guzman。用于自动驾驶的激光雷达：车载激光雷达与感知系统的原理、挑战和趋势。*IEEE*信号处理杂志，37(4): 50–61，2020。Zhaoyang Li、Yuan Wang、Wangkai Li、Rui Sun 和 Tianzhu Zhang。定位与扩展：一种用于点云少样本语义分割的解耦框架。载于欧洲计算机视觉会议，第18–34页。Springer，2024。Zhaoyang Li、Yuan Wang、Wangkai Li、Tianzhu Zhang 和 Xiang Liu。用于跨域少样本分割的双智能体优化框架。载于*IEEE/CVF*计算机视觉与模式识别会议论文集，2025d。Zhaoyang Li、Yuan Wang、Guoxin Xiong、Wangkai Li、Yuwen Pan 和 Tianzhu Zhang。通过大语言模型辅助超关系匹配实现广义少样本点云分割。载于*IEEE/CVF*国际计算机视觉会议论文集，2025e。

Hao Liu、Wilson Yan 和 Pieter Abbeel。语言量化自编码器：迈向无监督文本-图像对齐。神经信息处理系统进展，36：4382–4395，2023。

Zhipeng Luo、Zhongang Cai、Changqing Zhou、Gongjie Zhang、Haiyu Zhao、Shuai Yi、Shijian Lu、Hongsheng Li、Shanghang Zhang 和 Ziwei Liu。基于多层级一致性的无监督域自适应三维检测。载于*IEEE/CVF*计算机视觉国际会议论文集，第 8866–8875 页，2021。

Andres Milioto、Ignacio Vizzo、Jens Behley 和 Cyrill Stachniss。Rangenet++：快速且准确的激光雷达语义分割。载于*2019*年*IEEE/RSJ*智能机器人与系统国际会议（*IROS*），第 4213–4220 页。IEEE，2019。

Alexey Nekrasov、Jonas Schult、Or Litany、Bastian Leibe 和 Francis Engelmann。Mix3D：面向三维场景的上下文外数据增强。载于*2021*年国际三维视觉会议（*3DV*），第 116–125 页。IEEE，2021。

Yuwen Pan、Rui Sun、Wangkai Li 和 Tianzhu Zhang。探索恶劣条件下语义分割的天气感知聚合与自适应。载于*IEEE/CVF*国际计算机视觉会议论文集，2025。

Junsung Park、Kyungmin Kim 和 Hyunjung Shim。重新思考恶劣天气下鲁棒激光雷达语义分割的数据增强。*arXiv*预印本 *arXiv:2407.02286*，2024。

Junsung Park、Hwijeong Lee、Inha Kang 和 Hyunjung Shim。无物，即一无所有：突出恶劣天气下鲁棒激光雷达语义分割的安全关键类别。载于计算机视觉与模式识别会议论文集，第 6690–6699 页，2025。

Thierry Peynot、James Underwood 和 Steven Scheding。面向复杂条件下无人地面车辆的可靠感知。载于 *2009*年*IEEE/RSJ*国际智能机器人与系统会议，第1170–1176页。IEEE，2009年。

Gilles Puy、Alexandre Boulch 和 Renaud Marlet。使用华夫饼烤模进行汽车点云语义分割。载于 *IEEE/CVF*国际计算机视觉会议论文集，第3379–3389页，2023年。

Charles R Qi、Hao Su、Kaichun Mo 和 Leonidas J Guibas。PointNet：用于三维分类和分割的点集深度学习。载于 *IEEE*计算机视觉与模式识别会议论文集，第652–660页，2017年。

Ali Razavi、Aaron Van Den Oord 和 Oriol Vinyals。使用vq-vae-2生成多样的高保真图像。神经信息处理系统进展，第32卷，2019年。

Danilo Rezende 和 Shakir Mohamed。使用归一化流进行变分推断。载于国际机器学习会议，第 1530–1538 页。PMLR，2015。

David E Rumelhart、Geoffrey E Hinton 和 Ronald J Williams。通过反向传播误差学习表征。《自然》，323(6088)：533–536，1986。

Julian Ryde 和 Nick Hillier。激光和雷达测距设备在恶劣环境条件下的性能。《野外机器人学期刊》，26(9)：712–727，2009。

Christos Sakaridis、Dengxin Dai 和 Luc Van Gool。ACDC：用于语义驾驶场景理解的恶劣条件数据集及其对应关系。载于*IEEE/CVF*国际计算机视觉会议论文集，第 10765–10775 页，2021。

Rui Sun、Huayu Mai、Wangkai Li、Yujia Chen、Naisong Luo、Yuan Wang 和 Tianzhu Zhang。超越置信度：利用同质模式进行半监督语义分割。载于第四十二届国际机器学习大会，2025年a。

Rui Sun、Huayu Mai、Wangkai Li、Yujia Chen 和 Yuan Wang。两种损失，一个目标：为半监督语义分割对齐冲突梯度。载于*IEEE/CVF*国际计算机视觉会议论文集，2025年b。

Rui Sun、Huayu Mai、Wangkai Li 和 Tianzhu Zhang。迈向半监督语义分割中的无偏学习。载于第十三届国际学习表征会议，2025c。

Yuhta Takida、Yukara Ikemiya、Takashi Shibuya、Kazuki Shimada、Woosung Choi、Chieh-Hsin Lai、Naoki Murata、Toshimitsu Uesaka、Kengo Uchida、Wei-Hsiang Liao 等。Hq-vae：基于变分贝叶斯的分层离散表征学习。*arXiv* 预印本 *arXiv:2401.00365*，2023。

Haotian Tang、Zhijian Liu、Shengyu Zhao、Yujun Lin、Ji Lin、Hanrui Wang 和 Song Han。利用稀疏点体素卷积搜索高效的 3D 架构。载于欧洲计算机视觉会议，第 685–702 页。Springer，2020。

Hugues Thomas、Charles R Qi、Jean-Emmanuel Deschaud、Beatriz Marcotegui、Franc¸ois Goulette 和 Leonidas J Guibas。Kpconv：面向点云的灵活可变形卷积。载于*IEEE/CVF*国际计算机视觉会议论文集，第 6411–6420 页，2019。

Aaron Van Den Oord、Oriol Vinyals 等。神经离散表征学习。神经信息处理系统进展，30，2017。

Pascal Vincent、Hugo Larochelle、Yoshua Bengio 和 Pierre-Antoine Manzagol。利用去噪自编码器提取并组合鲁棒特征。载于第*25*届国际机器学习会议论文集，第 1096–1103 页，2008。

Pascal Vincent、Hugo Larochelle、Isabelle Lajoie、Yoshua Bengio、Pierre-Antoine Manzagol 和 L´eon Bottou。堆叠去噪自编码器：在深度网络中通过局部去噪准则学习有用的表征。机器学习研究期刊，11(12)，2010。

Yifei Wang、Zhengyang Geng、Feng Jiang、Chuming Li、Yisen Wang、Jiansheng Yang 和 Zhouchen Lin。用于多视图表征学习的残差松弛。神经信息处理系统进展，34:12104–12115，2021。

Aoran Xiao、Xiaofei Yang、Shijian Lu、Dayan Guan 和 Jiaxing Huang。FPS-Net：一种用于大规模激光雷达点云分割的卷积融合网络。*ISPRS* 摄影测量与遥感学报，176：237–249，2021。

Aoran Xiao、Jiaxing Huang、Dayan Guan、Kaiwen Cui、Shijian Lu 和 Ling Shao。PolarMix：一种适用于激光雷达点云的通用数据增强技术。神经信息处理系统进展，35：11035–11048，2022a。

Aoran Xiao、Jiaxing Huang、Dayan Guan、Fangneng Zhan 和 Shijian Lu。从合成数据到真实激光雷达点云的迁移学习，用于语义分割。载于*AAAI*人工智能会议论文集，第36卷，第2795–2803页，2022b。

Aoran Xiao、Jiaxing Huang、Weihao Xuan、Ruijie Ren、Kangcheng Liu、Dayan Guan、Abdulmotaleb El Saddik、Shijian Lu 和 Eric P Xing。自然环境中的三维语义分割：为恶劣条件点云学习泛化模型。载于*IEEE/CVF*计算机视觉与模式识别会议论文集，第9382–9392页，2023。

Jihan Yang、Shaoshuai Shi、Zhe Wang、Hongsheng Li 和 Xiaojuan Qi。ST3D：面向三维目标检测无监督领域自适应的自训练。载于*IEEE/CVF*计算机视觉与模式识别会议论文集，第 10368–10378 页，2021。

Jiahui Yu、Xin Li、Jing Yu Koh、Han Zhang、Ruoming Pang、James Qin、Alexander Ku、Yuanzhong Xu、Jason Baldridge 和 Yonghui Wu。通过改进的 VQGAN 实现向量量化图像建模。*arXiv* 预印本 *arXiv:2110.04627*，2021。

Jianlong Yuan、Yifan Liu、Chunhua Shen、Zhibin Wang 和 Hao Li。采用强数据增强的半监督语义分割简易基线。载于*IEEE/CVF*国际计算机视觉会议，第 8229–8238 页，2021。

Yang Zhang、Zixiang Zhou、Philip David、Xiangyu Yue、Zerong Xi、Boqing Gong 和 Hassan Foroosh。PolarNet：一种改进的网格表示，用于在线激光雷达点云语义分割。载于*IEEE/CVF*计算机视觉与模式识别会议论文集，第 9601–9610 页，2020。

Haimei Zhao、Jing Zhang、Zhuo Chen、Shanshan Zhao 和陶大程。Unimix：面向恶劣天气下域自适应且具有泛化能力的激光雷达语义分割。载于*IEEE/CVF*计算机视觉与模式识别会议论文集，第 14781–14791 页，2024。

赵恒爽、Li Jiang、贾佳亚、Philip H. S. Torr 和 Vladlen Koltun。点变换器。载于*IEEE/CVF*国际计算机视觉会议论文集，第 16259–16268 页，2021。

Lin Zhao、Hui Zhou、朱新革、Xiao Song、Hongsheng Li 和陶文兵。Lif-seg：用于三维激光雷达语义分割的激光雷达与相机图像融合。*IEEE*多媒体汇刊，第 26 卷：1158–1168，2023。

朱新革、Hui Zhou、Tai Wang、Fangzhou Hong、Yuexin Ma、Wei Li、Hongsheng Li 和林达华。用于激光雷达分割的圆柱形非对称三维卷积网络。发表于*IEEE/CVF*计算机视觉与模式识别会议论文集，第 9939–9948 页，2021 年。

附录 A

本附录提供更多细节和分析，以补充正文。我们介绍实现细节、消融研究以及更多可视化结果。具体而言，附录安排如下：

- **B 节** 实现：语义混淆先验（SCP）离散潜变量学习（VQ-VAE 设计、逐类量化、损失和重建）。
- **C 节**实现：语义偏移区域（SSR）定位与异常检测策略、掩码生成和膨胀。
- **D 节**讨论：离散编码与连续编码——为何离散表示更符合我们的先验和可解释性目标。
- **E 节**算法：A3Point 训练伪代码和端到端优化流程。
- **F 节** 增强：额外数据增强技术的影响分析与结论。
- **G 节**超参数：关键 A3Point 超参数（k, D, t, λ）的敏感性与权衡。
- **H 节**可视化：SSR 视觉分析及其与错误区域的关联。
- **I 节** 可视化：更多定性结果。
- **J 节** 评估：训练时 SSR 区域评估与推理时高失真子区域评估。
- **K 节** 训练开销：训练期间的额外成本及推理零开销说明。
- **L 节**对比：用原型（K-prototype）替代 VQ-VAE 的可行性及性能差距的来源。
- **M 节**计划：比较 50 轮与 15 轮训练计划及其影响。
- **N 节** 混淆：Things 类混淆矩阵及频率—增益关系分析。
- **O 节** 普遍性：不同数据集／架构中语义混淆的共性，以及在线先验的作用。
- **P 节**稀有类别：Oracle 在极稀有类别上出现交并比为 0.0 的现象，以及 A3Point 的稳定作用。
- **Q 节**EAS：在更大的增强空间下实现“安全训练”的能力及其上限。
- **第 R 节** 迭代：课程式增强上限与迭代式 A3Point。
- **第 S 节**公平性：为先前方法配备更强增强后的比较。
- **第 T 节** 码本：逐类 VQ-VAE 子码本的 t-SNE 可视化与解释。
- **第 U 节**方法比较：在相同增强空间下与 Mean Teacher、SCE 等的受控比较。
- **第 V 节**设计选择：消融实验验证为何在 SSR 中使用“全局最近代码”进行蒸馏监督。
- **第 W 节** 跨数据集的类别映射：统一的 19 类协议以及 [A]/[B]/[C] 的精确映射。
- **第 X 节** 影响：大语言模型使用说明和社会影响讨论。

B SCP 潜在学习的实现

本节将介绍语义混淆先验（SCP）潜在学习模块的实现细节。该模块基于向量量化变分自编码器（VQ-VAE）架构（Van Den Oord 等人，2017）。其目标是学习一种离散的潜在表示，以捕捉原始点云预测中的语义混淆模式。

B.1 向量量化变分自编码器架构

VQ-VAE 由编码器、解码器和向量量化层组成。编码器网络将输入的点云特征 x 与预测概率 $f(x)$拼接后，映射为潜在表示 z_e。我们在编码器中使用稀疏三维卷积（Tang 等人，2020）高效处理稀疏点云数据。架构如下：

- 编码器包含 4 个稀疏三维卷积下采样块，通道维度为 [16, 32,64, 128]。每个块由一个步长为 2 的稀疏三维卷积、批归一化和 LeakyReLU 组成。这会将点云下采样 16 倍。
- 随后再通过一个稀疏三维卷积、批归一化和 LeakyReLU，以及 2 个残差块，进一步处理这些特征。
- 最后通过稀疏三维卷积映射到潜在维度（64）。

解码器网络根据量化后的潜在表示 z_q 重建预测概率 $f(x)$。其架构与编码器镜像对称：

- 使用稀疏三维卷积、批归一化和 LeakyReLU，将潜在维度映射到初始特征维度（128）。
- 2残差块，用于进一步处理。
- 4 稀疏转置三维卷积上采样块，每个块由步幅为 2 的转置稀疏三维卷积、批归一化和 LeakyReLU 组成。这些块将特征上采样 16 倍，恢复到原始分辨率。
- 最后通过稀疏三维卷积和 tanh 激活重建预测概率。

B.2 按类别进行向量量化

为避免类间混淆，我们针对每个语义类别独立执行向量量化。VQ-VAE码本 $\mathcal{C}$ 包含每个类别的可学习嵌入，维度为 $C\times k\times D$，其中 C 是类别数， k 是每类嵌入数， D 是潜在维度。

在前向传播过程中，我们根据类别标签将编码器输出 z_e 拆分为类别特定潜变量。对于每个类别，使用 L2距离将潜变量量化为对应子码本中的最近嵌入。其他类别的嵌入距离会被屏蔽，以避免选错类别。

B.3 承诺损失和嵌入损失

学习目标结合了承诺损失和嵌入损失，遵循原始 VQ-VAE：

- 承诺损失促使编码器输出 z_e 趋向于所选嵌入 z_q，定义如下：

$$\mathcal{L}_{\text{commit}} = \|z_e - \text{sg}(z_q)\|_2^2 \tag{12}$$

- 嵌入损失使所选嵌入 z_q 接近编码器输出 z_e，其定义为：

$$\mathcal{L}_{\text{embed}} = \|\text{sg}(z_e) - z_q\|_2^2 \tag{13}$$

其中 sg 表示停止梯度操作。总向量量化损失为：

$$\mathcal{L}_{\text{VQ}} = \mathcal{L}_{\text{commit}} + \beta\mathcal{L}_{\text{embed}} \tag{14}$$

其中，β 为加权系数（设为 0.25）。

B.4 重建损失

为确保量化表示能够重建输入的预测概率，我们采用均方误差重建损失：

$$\mathcal{L}_{\text{recon}} = \|f(x) - \mathcal{D}(z_q)\|_2^2 \tag{15}$$

最终损失是重建损失与 VQ 损失的加权组合：

$$\mathcal{L} = \mathcal{L}_{\text{recon}} + \mathcal{L}_{\text{VQ}} \tag{16}$$

总而言之，我们的 SCP 潜在学习模块采用带有稀疏卷积的向量量化变分自编码器架构，以按类别的方式学习语义混淆模式的离散潜在表示。模型通过结合重建损失和 VQ 损失进行端到端训练。这样便能够从原始点云预测中捕获有关语义混淆的丰富先验信息。

C SSR 定位的实现

本节介绍语义偏移区域（SSR）定位模块的实现细节。该模块旨在识别增强点云中网络预测偏离已学习的语义混淆先验的区域，这表明数据增强可能导致了语义偏移。

C.1 跟踪潜在表示分布

在训练 SCP 潜变量学习模块期间，我们跟踪与每个量化潜变量（即码本中的每个嵌入）相关的潜在表示分布。具体而言，对于每个潜变量 e_i，我们计算其对应潜在表示在量化前的方差 σ_i^2。方差通过动量因子 γ 为 0.9 的指数移动平均（EMA）进行更新：

$$\sigma_i^2 \leftarrow \gamma\sigma_i^2 + (1-\gamma)\text{Var}(z_e|z_q = e_i) \tag{17}$$

其中，Var(·) 表示方差运算。这使我们能够估计每个潜变量对应的潜在表示的典型分布。

C.2 定位语义偏移区域

NS

为定位语义偏移区域，我们首先将增强点云输入训练好的编码器 $\mathbb{E}$，以获得其潜在表示 z_e。随后，我们将每个潜在表示与相应子码本中最近潜变量的分布进行比较。

具体而言，对于每个潜在表示 z_e^j，我们在相应类别的子码本中找到其最近潜变量$e_{NN(j)}$。如果 z_e^j 与$e_{NN(j)}$ 之间的距离超过基于所跟踪方差 $\sigma^2_{NN(j)}$设定的阈值，则将其视为离群点（即属于语义偏移区域）：

$$\text{isOutlier}(z_e^j) = \mathbb{1}\left(\|z_e^j - e_{NN(j)}\|_2 > t\sqrt{\sigma^2_{NN(j)}}\right) \tag{18}$$

其中 $\mathbb{1}(\cdot)$是指示函数，t 是用于控制阈值的超参数（设为 3）。其直观含义是，如果某个潜在表示明显偏离其最近潜变量的典型分布，那么它很可能对应语义偏移区域。

C.3 生成 SSR 和 SCR 掩码

识别出异常值潜在表示后，我们将其投影回点云空间，为语义偏移区域（SSR）和语义一致性区域（SCR）生成掩码。
我们首先在潜在表示空间中创建一个二值掩码 M_{outlier}，其中 M^j 离群点 =是 Outlier(z_e^j)。然后，我们使用点云下采样操作（在编码器中使用）的转置，将掩码上采样回原始点云的分辨率。由此得到 SSR 掩码 M_{SSR}。

SCR 掩码就是 SSR 掩码的补集：

$$M_{\text{SCR}} = \mathbb{1} - M_{\text{SSR}} \tag{19}$$

C.4 处理稀疏离群点

在实际应用中，异常值潜在表示可能稀疏且分散。为确保 SSR 掩码覆盖语义上有意义的区域，我们对 SSR 掩码执行膨胀操作。具体而言，对于 SSR 掩码中的每个点（x，y，z，c），我们将其一定半径 r 范围内的相邻点也设为属于 SSR。这有助于连接彼此靠近的离群点，并形成连续的语义偏移区域。
膨胀操作可以通过以下方式高效实现：首先将离群点掩码上采样至稠密三维网格，在网格上执行膨胀操作，然后使用最近邻插值将膨胀后的掩码下采样回点云。

总之，我们的 SSR 定位模块通过将潜在表示与学习到的语义混淆先验进行比较，在增强点云中识别语义偏移区域。该模块会生成 SSR 和语义一致性区域掩码，以便针对各个区域使用适当的损失来指导后续训练。该模块计算效率高，可以集成到训练流程中，而不会显著增加训练时间。

D 讨论：离散编码与连续编码

所提出的语义混淆先验（SCP）的一个关键方面是**显式获取表征的分布形式**，以便进行后续定位和优化。本节讨论离散编码方法与连续编码方法的区别，并解释为何离散编码更适合我们的框架。

D.1 编码范式比较

离散编码（例如 VQ-VAE）。VQ-VAE（Van Den Oord 等人，2017）通过以下方式学习离散潜在空间：1）通过编码器将输入映射到潜在空间。2）将潜在空间量化到码本中，从而有效地对代表性特征进行聚类。

连续编码方法。 连续表示方法可分为两类：**(1) 固定先验分布范式（例如，变分自编码器（Kingma 等，2013）、流模型（Rezende 和 Mohamed，2015；Kingma 和 Dhariwal，2018））**这些方法将潜在空间约束为固定的先验分布（例如，高斯分布），以实现隐式的全局分布对齐。然而，这种对齐缺乏可解释性，无法自然地解耦不同的语义类别，因此通常需要为每个类别分别配置编码器和解码器。**(2) 无约束潜在表示（例如，自编码器（Rumelhart 等，1986）、去噪自编码器（Vincent 等，2008）、堆叠去噪自编码器（Vincent 等，2010））**这些方法无需预定义先验即可学习连续潜在表示，因此难以直接提取结构化分布信息。因此，可能仍需采用聚类技术来施加结构，这与 VQ-VAE 的离散编码类似。

算法 1A3Point 的伪算法

输入： 源域 $D_S = \{(x_i^S, y_i^S)\}_{i=}^{N_S}$ 1
定义： 网络 f_θ、自动编码器 $\mathbb{E}$, $\mathbb{D}$、码本 $\mathcal{C}$、增强函数 $\mathcal{A}$、学习率 α, β, δ、损失权重 λ
输出： 训练后的模型 f_θ
对于每个批次 (x_i^S, y_i^S), 在 D_S **执行**
    # 步骤 ***1***：增强训练
    应用数据增强：$(\hat{x}_i^S, \hat{y}_i^S) = \mathcal{A}(x_i^S, y_i^S)$
    计算监督损失：$\mathcal{L}_{ce} = \ell_{ce}(f_\theta(x_i^S), y_i^S)$ 公式 (1)
    计算增强损失：$\hat{\mathcal{L}}_{ce} = \ell_{ce}(f_\theta(\hat{x}_i^S), \hat{y}_i^S)$ ▷ Eq. (2)
    # 步骤 ***2***：语义混淆先验学习
    计算潜在嵌入：$z_e = \mathbb{E}([f_\theta(x_i^S) \oplus x_i^S])$
    量化潜在编码：z_q 使用码本 $\mathcal{C}$ 量化 $= \text{quantize}(z_e)$ ▷ Eq. (3)
    重建预测：$\bar{x}_i^S = \mathbb{D}(z_q)$
    更新 VQ-VAE 损失：$\mathcal{L}_{vq} = ||x_i^S - \bar{x}_i^S||_2^2 + ||\text{sg}(z_e) - z_q||_2^2 + ||z_e - \text{sg}(z_q)||_2^2$ 式 (4)
    # 步骤 ***3***：语义偏移区域定位
    使用指数移动平均跟踪潜在编码 σ^2 的方差统计
    将增强预测映射到潜在空间：$z_e^{aug} = \mathbb{E}([f_\theta(\hat{x}_i^S) \oplus \hat{x}_i^S])$
    通过异常检测识别语义偏移区域（SSR）：
        - 语义一致性掩码：$M_{SCR} = \mathbb{1}(z_e^{aug} \in r(NN_{sub}(z_e^{aug})))$ 公式（7）
        - 语义偏移掩码：$M_{SSR} = \mathbb{1}(z_e^{aug} \notin r(NN_{sub}(z_e^{aug})))$ 公式（8）
    # 步骤***4***：优化
    计算 SCR 的损失：$\widetilde{\mathcal{L}}_{ce} = \ell_{ce}(f_\theta(\hat{x}_i^S) \odot M_{SCR}, y_i^S \odot M_{SCR})$ 公式（9）
    计算 SSR 的蒸馏损失：$\mathcal{L}_{distill} = ||z_e^{aug} - \text{sg}[NN_{global}(z_e^{aug})]||_2^2$ ▷ 公式（10）
    更新模型：$\theta \leftarrow \theta - \delta\nabla_\theta(\mathcal{L}_{ce} + \mathcal{L}_{ce} + \lambda\mathcal{L}_{distill})$ 公式（11）
结束循环

D.2 为什么离散编码是必要的?

尽管离散编码将潜在模式的数量限制为 k，但这并不影响我们框架的有效性。相反，它带来了若干优势：

更好的可解释性： 每个潜在编码代表一种特定的语义混淆模式，因此更容易分析和解释所学习到的表征。

Class 维分离： VQ-VAE 能够自然地按类别对潜在表示进行聚类，
对 至关重
要，…尤其是对我们的语义偏移定位而言。

鲁棒性与泛化： 精心设计的码本可确保模型仅学习最具代表性且最有意义的模式，从而提升对未见数据的泛化能力。

E OFA3POINT伪算法。

我们还在算法1中提供了详细的算法描述。

F 其他增强技术的影响

To f为进一步分析各种增强技术的作用，我们开展实验来评估 e
thei其对源域和目标域性能的影响。

F.1 实验设置

我们遵循第3.3.2节中的相同训练设置，调整缩放、翻转和噪声扰动等次要数据增强的强度，同时保持点变形和点损失增强不变。我们定义了三个增强级别：- **轻度**：采用主要实验中使用的默认增强设置。- **中度**：增大缩放和翻

转的幅度，同时保持原始分布。- **重度**：大幅增大缩放因子和噪声扰动级别。

F.2 结果与分析

表 7 展示了 [A]→[C]和 [B]→[C] 基准上的结果。

增强级别	[A]→[C]	[B]→[C]	[A]→[A]
轻度	**41.3**	**27.2**	**62.5**
中度	40.1	26.0	60.3
重度	37.5	24.5	55.8

表 7：不同增强级别对领域泛化性能的影响。

从结果中，我们观察到：1. **过度的次级增强会降低源域和目标域性能**。将增强幅度从轻度提高到重度后， [A]→[C] 上的平均交并比降低了 3.8%， [B]→[C]上的平均交并比降低了 2.7%。2. **中度增强带来的收益有限**。与轻度设置相比，中度增强会略微降低性能，表明次级增强对领域自适应的贡献并不显著。3. **随着增强强度的增加，源域性能会下降**。在 [A]→[A],上，平均交并比从 62.6% 显著下降至 55.8%，这表明过度的缩放和噪声扰动会破坏原始数据分布，并阻碍学习。

这些发现表明，点变形和损失（抖动、点丢弃）是领域自适应的主要驱动因素，而其他增强实际上属于领域无关因素，应谨慎应用，以避免性能下降。

G 参数设置的影响

本节讨论 A3Point 框架的关键超参数及其对模型性能和行为的影响。所有实验均在 SemanticKITTI→semanticSTF 上进行。

G.1 潜在空间维度

SCP 潜变量学习模块中潜在空间的维度由每类嵌入数量（k）和嵌入维度（D）决定，它们会影响所学习语义混淆模式的表达能力和细粒度程度。

k	D	平均交并比（%）
16	64	39.9
32	64	41.3
64	64	41.4
32	32	40.8
32	64	41.3
32	128	41.2

表 8：潜在空间维度的消融研究。

在我们的实现中，我们设置了 $k = 32$ 和 $D = 64$ （表 8）。这些取值在潜在空间的表征能力与模块的计算效率之间取得了良好的平衡。增大 k 可使模型捕获每个类别中更多样的语义混淆模式，但也会增加内存占用和过拟合风险。同样，增大 D 能提升每个嵌入编码更复杂模式的能力，但也会增加计算开销。

G.2 SSR 定位阈值

SSR 定位模块中使用的阈值 t 决定了检测语义偏移区域的灵敏度。

	$t=2$	$t=3$	$t=4$
平均交并比 (%)	41.0	41.3	40.9

表 9：SSR 定位阈值的消融研究。

在我们的实验中，我们设置 $t=3$，这能有效识别语义偏移显著的区域，同时最大限度地减少假阳性。较高的 t 会采取更保守的策略，只有与已学习的语义混淆模式存在最显著偏差的区域才会被视为语义偏移。相反，较低的 t 会使模块更加敏感，可能将更多区域识别为语义偏移。

G.3 损失加权系数

损失加权系数 λ 平衡了交叉熵损失和蒸馏损失在整体训练目标中的贡献。

	$\lambda=0.02$	$\lambda=0.1$	$\lambda=0.5$
平均交并比 (%)	40.5	41.3	40.9

表 10：损失加权系数的消融研究。

在实验中，我们发现设置为 $\lambda=0.1$ 可以很好地平衡这两项损失。较高的 λ 会更加重视蒸馏损失，促使模型更专注于使语义偏移区域中的预测与学到的语义混淆模式保持一致。另一方面，较低的 λ则会更加重视交叉熵损失，优先保证整体分割准确率。

H SSR 的视觉分析

为了更好地理解我们的方法如何识别语义偏移区域（SSR），我们进行了详细的视觉分析，对比误差图与检测到的 SSR，如图 6 所示。误差图突出显示模型预测与真实标签不同的区域，而 SSR 则表示由我们的语义偏移检测机制识别出的区域。

从可视化结果中，我们观察到 SSR 与容易出现预测错误的区域之间存在很强的相关性。具体而言：（1）我们的 SSR 检测能够有效捕捉恶劣天气条件导致显著语义歧义的区域，尤其是物体边界处以及点云密度降低的远距离区域。（2）SSR 通常对应于具有挑战性的场景，例如不同语义类别之间的交界处（如道路与人行道的边界）以及几何结构复杂的区域。（3）语义一致性区域（非 SSR 区域）通常与模型能够保持准确预测的区域高度吻合，验证了我们的方法识别可靠预测的能力。这一视觉分析表明，我们的 SSR 检测机制能够为训练期间采用不同的优化策略提供有效指导。

I 更多定性结果

为进一步展示我们所提出方法的有效性，我们在图 7 中给出了更多定性结果，将我们的方法与基线方法在具有挑战性的 Se-manticKITTI→SemanticSTF 域泛化任务上进行比较。这些定性结果进一步验证了语义混淆先验学习和语义偏移区域定位策略在提升激光雷达语义分割的域泛化性能方面的有效性。

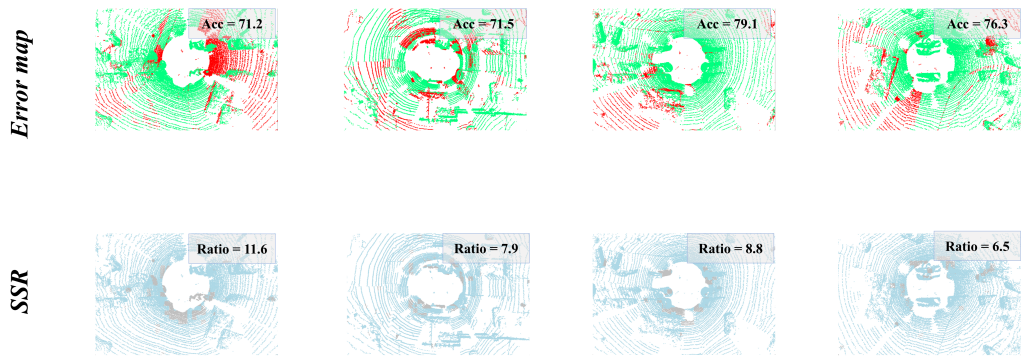


图 6：误差图与语义偏移区域对比的定性结果。

J SSR 和高失真区域中的针对性评估

附录 H 展示了 SSR 与误差图之间的相关性，可见二者与易出错区域有大量重叠。为更直接地量化改进效果，我们增加了两项互补评估：

- **训练时 SSR 区域评估（近似平均交并比）：** 我们使用训练时的样本及其 SSR 掩码生成固定的评估掩码，并跟踪 SSR 区域内平均交并比的变化：

  – 使用相同规则生成固定评估掩码 $\mathcal{M}_{\text{SSR eval}}$（我们采用仅 EAS 设置下获得的先验编码器和方差统计）。– 由于 SSR 中的真实标签并不可靠（根据定义，经过强数据增强后标签会发生错位），我们遵循半监督实践，维护一个指数移动平均（均值教师）模型作为代理评估器。这里指数移动平均模型优于真实标签，因为它能提供稳定、去噪的目标，反映模型的长期判断，同时不会继承强数据增强引入的语义错位。在 SSR 区域中，与指数移动平均模型的一致性更高，表明模型恰恰在真实标签噪声较大的区域具有更好的稳定性和可靠性。

  如图 8 所示，与仅使用 EAS 相比，A3Point 在 $\mathcal{M}_{\text{SSR eval}}$ 中的预测与 EMA 教师的吻合度更高。具体而言，加入 SSR 蒸馏后，SSR 区域内的预测与教师的结果更加一致。

- **推理时高失真子区域评估：** 如图 9 所示，我们使用密度/曲率阈值近似确定“高失真”子区域。在这些子区域中，A3Point 在出现频率最高的五个类别上提升了 1.9–15.1 个百分点，平均交并比提高了 7.6 个百分点，优于 EAS。构建细节：

  – 对于每个点 i，使用 $k = 32$ 查找 k 个最近邻 $N_k(i)$。– 局部密度(i) $\approx$ 到第 k 个最近邻距离的倒数。– 通过对 $N_k(i)$ 进行 PCA 计算局部曲率：计算协方差矩阵 Σ_i，得到特征值 $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$，并定义曲率$(i)$ $= \lambda_3/(\lambda_1 + \lambda_2 + \lambda_3)$。– 通过取均值，将逐点密度/曲率聚合到体素 v，得到密度(v) 和曲率(v)。– 使用基于分位数的阈值定义高失真指示函数（选取最困难的 20% 区域）：HighDist(i) $= \mathbb{I}$(密度(i) $\leq \tau_d \vee$ 曲率(i) $\geq \tau_c$)。– 如表 11 所示，在这些高失真掩码内，A3Point 的表现优于 EAS。

这些分析表明，我们的方法不仅提升了整体指标，还切实增强了训练期间检测到的高风险区域内的性能。

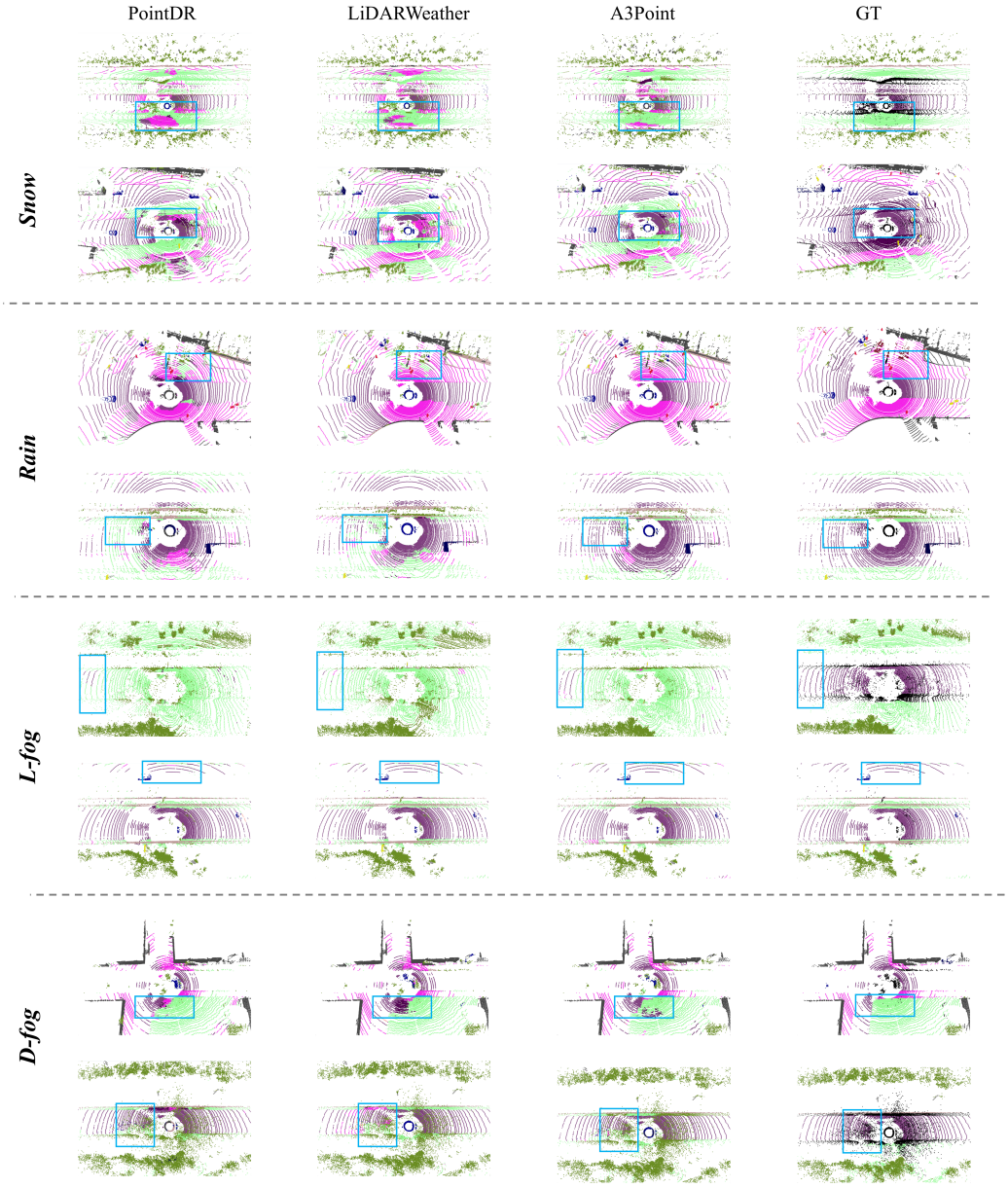


图 7： [A] → [C], 上的定性结果，改进之处以方框标出。

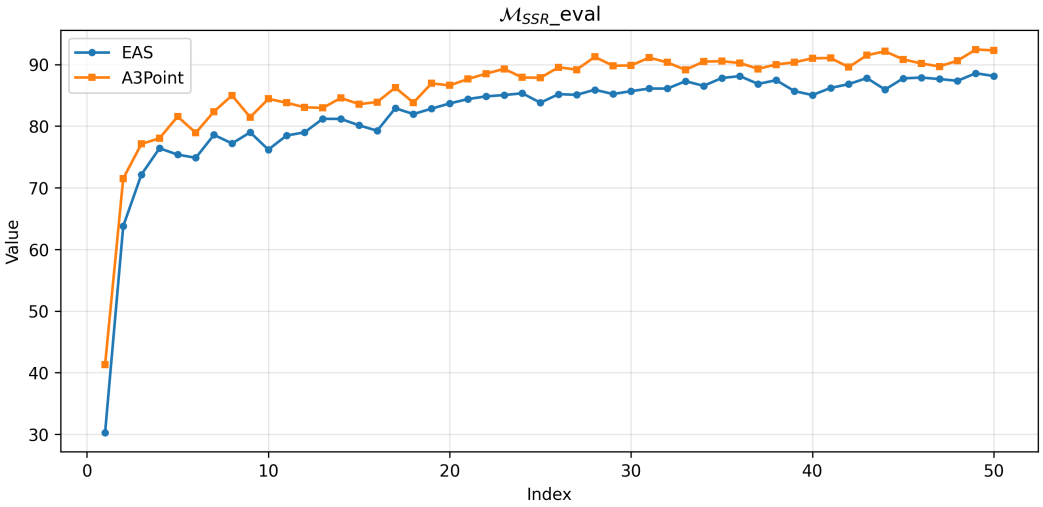


图 8：与 EMA 教师在 $\mathcal{M}_{\text{SSR 评估}}$中的一致性。

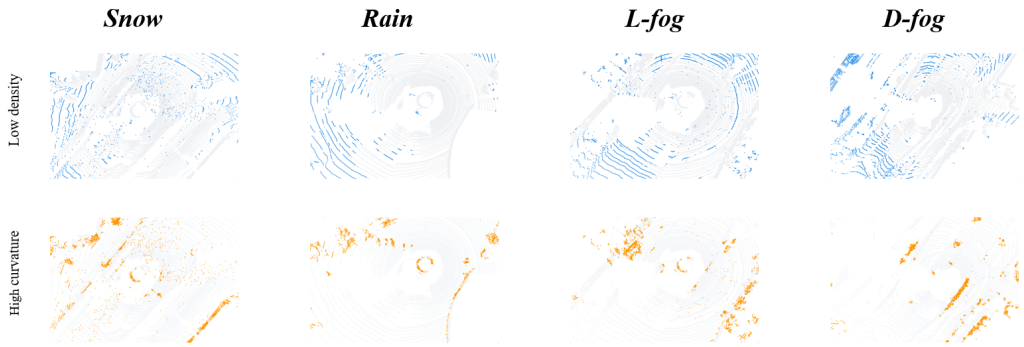


图 9：展示“高畸变”子区域的示例场景。

表 11：高失真子区域内的性能。

方法	Car	Road	人行道	植被	地形	总体平均交并比
EAS	61.6	50.7	21.4	47.5	35.7	27.1
A3Point	76.7	57.6	26.8	51.5	37.6	34.7

K 训练时间开销

A3Point 引入了一个轻量级的按类别划分的 VQ-VAE 模块——该模块采用稀疏卷积，编码维度为 $D = 64$，每类码本大小为 $k = 32$（架构详情见附录 B.1）——以便在训练期间实现 SCP 和 SSR。下面详细说明其实际影响和成本：

- **无推理时间开销。** SCP 和 SSR 都是仅用于训练的组件。它们在测试时会被禁用，因此推理延迟和内存占用与基线模型完全相同。导出的检查点和前向路径均保持不变。
- **通过指数移动平均低成本地统计。** 我们使用指数移动平均（EMA）维护逐码字统计量（方差）。每次迭代仅需进行 $O(1)$次更新，额外内存开销可忽略不计，且不会随数据集大小或批次长度增加。
- **用于 SSR 定位的冻结先验。** 用于定位 SSR 区域的编码器在定位过程中保持冻结，且不进行反向传播。此设计将额外计算量限制在轻量级前向传播范围内，并避免梯度开销，从而保持利用率稳定。

为清晰展示开销，我们在相同硬件和训练设置下报告一个具有代表性的运行时快照，见表 12。鉴于 SemanticSTF 上显著的平均交并比增益（相较基线，在 [A]→[C]and [B]→[C] 上分别提升 +9.9 / +11.7），我们认为训练时间的增加是一种值得的权衡：将更多预算集中在训练端，以获得更强的恶劣天气鲁棒性，同时保持推理路径简洁且不变。

方法	**GPU显存（MB）**	**每秒迭代次数（it/s）**	**总时间（h）**
基线	8,496	2.72	24.4
PointDR	9,624	1.97	33.7
EAS	8,932	2.05	32.4
A3Point（本文）	17,144	1.61	41.3

表12：相同硬件和设置下的训练时间开销对比。

L 论将VQ-VAE替换为基于原型的替代方案

我们评估了一种简单的基于原型的方法能否替代 VQ-VAE先验。具体而言，我们实现了一个类别级K原型变体（$k = 32$），作为 VQ-VAE模块的即插即用替代方案。在训练预算和数据增强设置相同的情况下，VQ-VAE 的表现始终更好，如表13所示。

设置	**[A]→[C]**	**[B]→[C]**
None	31.4	15.5
EAS	38.7	24.9
基于原型	40.0	25.4
基于VQ-VAE（本文）	**41.3**	**27.2**

表13：在相同设置下，逐类K-prototype与VQ-VAE先验的比较 gs.

我们将这种差距归因于 VQ-VAE先验的三项优势：

- **类内多模态性**。 类内的错误/歧义模式本质上具有多模态性（例如，路缘附近的车道线标记、部分遮挡、远距离稀疏性）。VQ-VAE 学习到的逐类码本比单个原型质心更忠实地捕捉多种模式，从而能更准确地定位高风险子区域。
- **规范的异常建模**。 VQ-VAE 支持跟踪每个码字的方差（通过指数移动平均），从而生成紧凑且近似服从高斯分布的局部簇，并为异常检测提供可解释的“先验半径”。原型方法通常需要额外的度量学习损失或经过精细调优的阈值才能良好运行，而且在不同条件下可能表现脆弱。
- **在线自适应性与稳定性**。 我们的先验会随着骨干网络的演变同步在线更新，同时 SSR 定位编码器在反向传播期间保持冻结，从而提供稳定而响应迅速的统计信息。随着特征空间发生变化，原型更新可能滞后或漂移，导致训练过程中 SSR 定位质量下降。

简言之，VQ-VAE 提供了一种更稳健、可解释且经实证效果更强的逐类先验建模机制，只需略微增加训练成本，便能带来稳定的性能提升。

M 训练计划：50 个训练轮次 vs. LiDARWeather 的 15 个训练轮次

所有主要结果和消融实验（表 3–6）均采用类似 PointDR 的 50 轮训练计划，以确保内部比较公平。为区分训练计划长度与算法带来的提升，我们分别采用短训练计划（15 轮，LiDARWeather 风格）和长训练计划（50 轮）进行受控比较。

短训练计划（15 个训练轮次）在 15 轮训练计划下（表 14），A3Point 仍比我们复现的 LiDAR Weather 高出 $+2.7$–4.5 个平均交并比点，且训练预算相同，这表明我们的性能提升并不依赖于更长的训练时间。

方法	[甲]→[丙]	[乙]→[丙]
基线	31.4	15.5
LiDARWeather*	37.8	20.4
A3Point（我们的方法）	**40.5**	**25.9**

表 14：15 轮训练（LiDARWeather 风格）。* 表示复现结果。

长训练计划（50 轮）将基线和 LiDARWeather 的训练延长至 50 轮（表 15）仅带来了有限的提升，且两者仍逊于 A3Point，这表明仅延长训练计划并不能弥合差距。

方法	[A]→[C]	[B]→[C]
基线*	32.3	16.1
LiDARWeather*	38.6	21.6
A3Point（我们的方法）	**41.3**	**27.2**

表 15：训练 50 个周期。* 表示复现结果。

N “物体”类别的混淆矩阵

我们在图10中列出了 Oracle 和 A3Point 在 Seman- ticSTF 上的混淆矩阵及各类别的点数，以阐明类别频率与性能提升之间的关系。我们的主要观察如下：

- **频率并不足以预测性能提升**。诸如行人和骑摩托车者等类别，尽管总体出现频率不低，却仍常常受到严重遮挡，

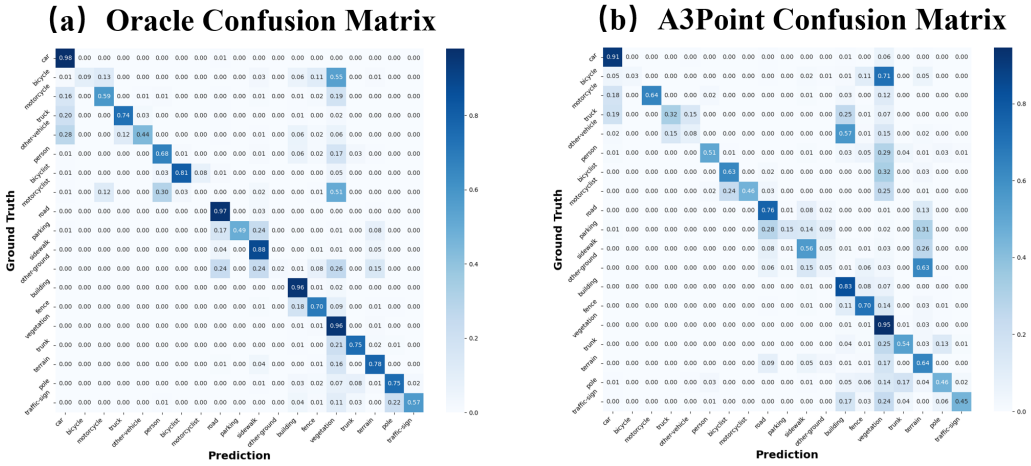


图 10：SemanticSTF 上的混淆矩阵。左：Oracle。右：A3Point。

并存在姿态/形状变异性和远距离稀疏性。在恶劣天气下，这些因素会加剧类间混淆（例如，行人与杆/标志，骑摩托车者与自行车/摩托车之间的混淆），从而减少每个场景中的可靠点数量，并限制可实现的改进幅度。

- **几何特征鲜明且稳定的类别增益更大**。例如自行车、摩托车、交通标志和汽车等类别，在扰动下仍能保留更清晰的几何特征（例如边缘、平面以及可重复的部件构型）。我们的 SCP/SSR 先验有助于保持这些区域的一致性，从而带来更显著的提升，这与表 1–2 中报告的趋势一致。

总体而言，混淆矩阵表明，鲁棒性提升更能由天气变化下局部几何线索的稳定性和显著性来解释，而非仅由原始类别频率决定。

语义混淆的普遍性

Thi本节阐明“语义混淆”如何在不同数据集和架构中体现 h
thre互补点：

1. **数据集驱动的模式与模型驱动的强度**。在不同架构之间，哪些类别容易混淆，很大程度上取决于数据集几何、密度和采样统计特性（例如道路–人行道、汽车–卡车）。相比之下，混淆的程度（置信度分布、错误率）更多取决于模型的容量、正则化和训练动态。在我们的设置中，SCP 会针对每个骨干网络，在源域上在线学习混淆先验；对于固定的骨干网络，混淆模式保持稳定，而其强度会随着主网络性能提升而减弱。

2. **通过在线先验实现架构对齐**。A3Point 从当前模型在原始域上的预测中学习按类别划分的离散先验（通过 VQ-VAE），并将其冻结，用于对增强样本进行异常检测。这种“在线拟合、冻结检测”的设计使先验与模型不断演变的决策边界保持一致，并避免跨架构不匹配。

3. **跨骨干网络和设置的实验证据**。表3显示，在两个体素骨干网络（MinkowskiNet-18/32w 和 SPVCNN）以及三种域泛化轨迹（[A]→[C]，[B]→[C]，[A]→[D]）上，均取得了稳定的提升。尽管绝对混淆程度会随骨干网络而异，A3Point 始终能提升鲁棒性。

此外，我们还评估了多模型（离线）先验：将三个独立训练的源域模型的混淆统计信息进行聚合，并在训练期间冻结该先验。如表16所示，在 [A]→[C] 上，平均交并比从无先验时的 38.7 提升至使用离线先验时的 40.7，以及使用基于集成的离线先验时的 41.0；在 [B]→[C], 上，平均交并比则分别从 24.9 提升至 26.4 和 26.9。尽管如此，这两种变体的表现仍逊于我们的在线先验——该先验随当前模型的决策边界共同演化——其在 [A]→[C] 上达到 41.3，在 [B]→[C] 上达到 27.2。我们认为，这一差距源于根本性的不匹配：离线先验是静态的，无法跟踪训练期间不断演变的决策边界、校准和特征几何结构（尤其是在强数据增强下）；而我们的在线先验会在原始域上更新，仅在对增强数据进行检测时冻结，从而在整个训练过程中保持对齐。这些发现与正文中的分析一致（第456–467行；另见表6）。

策略	**[A]→[C]**	**[B]→[C]**
None	38.7	24.9
离线	40.7	26.4
离线（集成）	41.0	26.9
在线（我们的方法）	**41.3**	**27.2**

表16：离线和在线混淆先验的比较。在线（我们的方法）通过与不断变化的决策边界保持一致，始终取得最佳平均交并比。

P 稀有类别上的 Oracle 0.0 交并比与 A3Point 改进

重新审视 Oracle 结果： [C], 观察到的 摩托车/骑摩托车者 交并比为 0.0，是极端类别不平衡和跨天气变化下可复现性不稳定所致。多次相互独立的重跑得到非零交并比，且方差显著，这表明 0.0 是方差伪象，而非系统性失效。下方展示了五次独立运行的结果。

Run	car	bi.cle	mt.cle	truck	oth-v.	pers.	bi.clst	mt.clst	road	parki.	sidew.	other-g.	build.	fence	veget.	trunk	terr.	pole	traf.	**mIoU**
1	89.4	42.1	0.0	59.9	61.2	69.6	39.0	0.0	82.2	21.5	58.2	45.6	86.1	63.6	80.2	52.0	77.6	50.1	61.7	54.7
2	89.6	37.4	9.6	56.4	64.3	67.3	44.8	0.8	81.8	27.9	54.6	50.1	86.5	66.2	78.3	56.3	75.4	45.8	58.6	55.3
3	89.7	41.8	4.0	58.6	62.8	66.1	40.7	0.0	82.0	23.4	51.8	47.2	86.7	64.9	78.7	59.1	75.0	49.0	57.0	54.6
4	89.6	35.3	0.2	60.8	64.3	67.4	44.4	5.4	81.3	27.4	54.6	50.2	87.2	66.3	77.8	61.1	74.0	51.8	59.5	55.7
5	89.7	39.9	6.8	58.1	62.1	61.6	48.5	11.1	76.4	21.7	51.4	43.5	83.1	61.3	77.2	63.4	74.9	55.3	56.4	54.9

表17： [C] 上的 Oracle：五次独立运行中各类别的交并比。稀有类别（摩托车、骑摩托车者）表现出较高方差；有些运行中交并比为 0.0，另一些则没有。

这种不稳定性也解释了为什么其他方法（包括基线）在摩托车/骑摩托车者上也有明显提升：在源域 [A] 上训练能提供稳定且数量充足的样本，缓解 [C] 上样本稀缺的问题。这种效果由数据集驱动，普遍可见，并非某一种算法所独有。

除数据集覆盖范围外，A3Point 还包含两个组件，尤其有助于在强扰动下处理长尾小尺寸类别：

- **借助 EAS 扩大曝光范围**。增强空间使点密度和姿态更加多样，从而提高脆弱类别的召回率；这些类别在远距离处往往会受到稀疏性和遮挡的影响。
- **借助 SSR 实现更安全的监督**。当数据增强引发语义偏移时，SSR 会以潜在空间蒸馏取代硬监督，使其向最佳匹配的代码靠拢。这种基于先验的一致性指导减少了标签–观测不匹配，并使稀有类别的训练更加稳定。

实际上，相较于朴素的 Oracle，重增强基线已经提升了稀有类别的表现；SSR 的潜在引导进一步减少了假阳性和有害更新。总体效果是带来更稳定的

摩托车/骑摩托车者 等稀有类别的性能提升。修订后的文本阐明了数据集带来的改进与 A3Point 特有的稳定作用之间的区别。

Q 安全允许更大增强空间的有效性

我们补充实验，以直接评估 A3Point 是否能够安全地扩展增强空间。

- **默认 EAS 的范围**。默认范围已覆盖常见街景 LiDAR 退化。超出这些界限（例如极端点丢弃比例或抖动）很快会变得在物理上不合理，并带来标签–几何不匹配的风险，从而可能破坏监督。
- **任意大范围的压力测试**。 我们将抖动标准差扩展至 $[0, 0.10]$ ，将点丢弃比例扩展至 $[0, 0.99]$ （无–过度），保持所有其他数据增强不变，并在 [A]→[C]上进行评估。

方法 无 轻度 中度 重度 随机（轻度–重度） 过度 随机*（无–过度）					
基线 31.4 37.6	38.0	37.1	38.7	5.1	36.5
A3Point 31.8 38.5	40.4	40.5	**41.3**	7.8	**41.2**

表18：在 [A]→[C] 上进行扩大数据增强范围的压力测试。随机：在轻度至重度范围内采样；随机*：在无至过度范围内采样。

如表 18 所示，我们有以下关键观察：
– 不使用 SSR 的过度数据增强有害。在范围极大的情况下，基线性能严重下降，这与语义偏移导致的错误监督一致。– SSR 可在重度失真下稳定训练。
A3Point 在无→过度范围内随机采样时仍能保持良好性能，因为 SSR 会定位偏移区域，并切换到潜在空间蒸馏以避免有害更新。– 物理规律设定了一个硬性上限。当丢弃率 $\to 1$ 和抖动标准差 $\to 0.1$时，输入不再像合理的激光雷达扫描；所有方法的性能均会下降，这表明了有效数据增强的实际极限。

- **自适应行为的证据**。 如图5所示，SSR掩码比例随着数据增强强度的增加而上升，表明检测器会自适应地将更多区域标记为偏移区域，且偏移程度越大，标记越多。这种自适应性使得安全使用更强的数据增强成为可能：与学得的混淆先验一致的区域接受标准监督，而偏移区域则接受与先验一致的软指导。

总而言之，A3Point 通过缓解语义偏移发生处的错误监督，扩大了可用的增强范围，同时遵循一个基于物理的上限，超过该上限后，增强将不再有益。

R 扩展：具有可变增强空间的迭代式 A3Point

迭代式优化混淆先验（例如混淆矩阵或 SCP）可以进一步提升鲁棒性和泛化能力。纯离线先验的主要局限在于其静态性质：与 Q1 中的多模型集成类似，它们无法跟踪当前网络不断演变的决策边界、校准状态和容量。在实践中，这类静态先验往往与模型的置信度分布和错误模式不匹配——尤其是在训练早期——从而导致其结果逊于在线演化。

相比之下，我们的方法会在原始域上在线更新混淆先验，仅在对增强输入进行异常检测时将其冻结。这种设计不要求先验完美；

相反，它在整个训练过程中始终与模型的当前状态保持一致。实证结果表明，与离线混淆矩阵相比，自适应在线先验带来了更好的稳定性和更高的最终平均交并比（见表6），尤其显著提升了长尾类别和长距离稀疏区域的表现。

为评估课程式扩展，我们实现了 A3Point 的迭代变体，分阶段扩展增强空间。训练分为三个阶段：从轻度增强上限开始，训练若干轮次；使用由此得到的 SCP 作为初始化，并将上限提高到中度；最后，将上限提高到重度。因此，增强上限依次为轻度 →中度 → 重度，且每个阶段都使用上一阶段学得的 SCP 进行初始化。结果汇总如下。

策略	**[A]→[C]**	**[B]→[C]**
None	38.7	24.9
A3Point	41.3	27.2
A3Point+ （课程式）	**41.5**	**27.8**

表 19：分阶段扩展增强上限的迭代式课程风格 A3Point（轻度→ 中度 → 重度）。

如表 19 所示，扩展到更大的增强空间时，迭代方案带来的额外收益有限。这是一个颇有前景的方向——利用 SCP 统计数据，在闭环调度中自适应调整增强分布和 SSR 阈值——不过，主要的改进仍源于保持在线对齐，而非离线聚合。

为先前方法采用更强的增强

为确保公平性，我们使用与增强型增强空间（EAS）相同的上限重新运行先前的基于增强的方法：点丢弃率在 $[0.2, 0.8]$ 中采样，抖动在 σ 至$[0.01, 0.05]$之间采样，其他方面则保持其原始流程不变。

方法	**[甲]→[丙]**	**[乙]→[丙]**
基线	31.4	15.5
基线 +EAS	38.7	24.9
PointDR*	33.9	19.8
PointDR* +EAS	39.1	25.1
LiDARWeather*	37.8	20.4
LiDARWeather* +EAS	39.6	25.9
A3Point（我们的方法）	**41.3**	**27.2**

表20：在相同的 EAS 上限下重新运行先前方法。* 表示复现结果。

在更强的数据增强下，先前方法的性能确实有所提升，但增益逐渐饱和，且仍不及 A3Point。仅扩展增强空间并不足够，除非明确处理语义偏移；SSR 定位和潜在蒸馏使更强的数据增强真正变得可用。

T 可视化学习到的码本

我们使用 t-SNE 可视化学习到的 VQ-VAE 码本，如图11所示。每个类别的子码本以不同颜色呈现；标记大小表示通过指数移动平均跟踪的每码方差，让人一目了然地了解原型稳定性和覆盖范围。

可视化中观察到的一致模式：

- **类内多模态性**。 在每个类别内，原型形成多个簇，分别对应不同的局部几何结构和上下文配置（例如，道路–人行道边界、汽车–植被遮挡边缘、地形不连续处）。

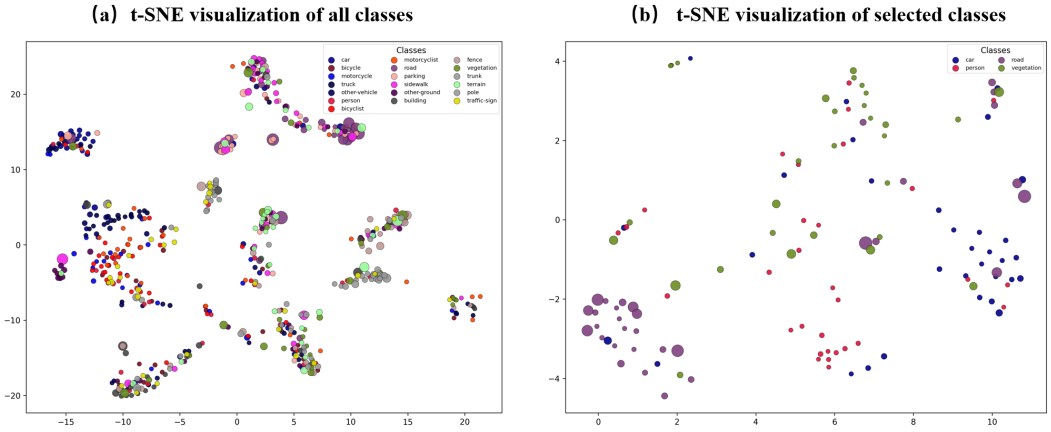


图 11：学习到的 VQ-VAE 码本的 t-SNE。颜色表示类别专属子码本；标记大小编码每码方差。

码本捕捉细粒度、反复出现的混淆模式，而不是将其压缩为单一类别模板。

- **类间邻域。** 来自语义相似或常被混淆类别的原型（例如，道路 与 人行道，建筑物 与围栏）彼此邻近，反映了分割网络所遇到的典型混淆流形。
- **渐进式专门化。** 在训练过程中，原型沿特定于混淆的轴变得更加紧凑且分离得更好，表明专门化程度不断提高。这种锐化改善了针对语义偏移的异常检测：偏离成形良好的混淆簇的嵌入更容易被识别为离群点。

通过在潜在空间中显式编码混淆流形，码本为区分真实的语义混淆（预期的、类别一致的变异性）与增强诱导的语义偏移（标签–几何不匹配）提供了原则性基础。它并非记忆类别均值，而是组织了一幅混淆模式的离散图谱，SSR 定位可借此判断预测何时一致、何时发生偏移。

U 与其他领域泛化方法的直接比较。

我们在与 A3Point 相同的增强空间内进行受控比较。具体如下：

- **基于数据增强的 Mean Teacher (MT)：** 学生模型接收强增强输入，而教师接收原始或弱增强输入；采用 KL 散度一致性损失。
- **对称交叉熵（SCE）：** 一种对标签噪声具有鲁棒性的目标函数。

方法	[甲]→[丙]	[乙]→[丙]
基线	31.4	15.5
EAS	38.7	24.9
EAS + MT	39.0	25.5
EAS + SCE	38.8	24.6
A3Point（我们的方法）	**41.3**	**27.2**

表21：在相同增强空间下与其他领域泛化方法的比较。

如表21所示，我们有以下观察：

• MT 与 SCE 仅带来有限提升，但并未显式识别或处理强扰动下由增强引起的语义偏移，从而导致对未对齐伪目标的过拟合，或不得不采用较弱的数据增强。

• A3Point 结合了 (i) 从未增强预测中学习的按类别的语义混淆先验，(ii) 用于定位语义偏移区域（SSR）的潜在空间异常检测，以及 (iii) 区域自适应监督，从而在恶劣天气目标上实现更稳定的性能提升和更高的平均交并比。

V 为什么 SSR 要采用全局最近编码?

动机。强增强可能会在 SSR 期间引入标签与语义错位。使用类别条件最近编码会将监督锚定到原始类别码本，从而可能强化错误标签带来的偏差。相比之下，全局最近编码提供了一种与类别无关、以稳定性为导向的先验，能在潜在空间中将偏移区域与最兼容的混淆模式对齐，而不会受原始类别（可能是错误类别）的牵制。

变体	**[A]→[C]**	**[B]→[C]**
基线	31.4	15.5
EAS	38.7	24.9
类别条件最近编码（SSR）	38.8	24.5
基于带温度的 logits 的软交叉熵（SSR）	39.7	25.8
与增强前预测保持一致	37.2	24.4
A3Point（潜在空间中的全局最近代码）	**41.3**	**27.2**

表22：在相同增强空间下对SSR监督方案的受控比较。全局最近代码（类别无关的潜在对齐）表现最佳。

解读。 我们在表22中进行受控比较：

• **类别条件最近编码**。 将监督绑定到原始的类别特定码本。它无法解决强增强下的标签–语义不匹配，并且仍然容易受到原始类别偏差的影响，相对于EAS仅带来可忽略的提升。

• **对logits使用带温度的软交叉熵**。 温度缩放的软交叉熵降低了过度自信，并部分缓解了语义偏移，带来轻微增益。然而，它仍然以标签为条件，无法完全纠正不匹配。

• **增强前一致性**。 当数据增强引发语义偏移时，将增强前预测用作教师容易受到影响；强制这种一致性会强化不可靠的对齐目标，从而使性能相对于EAS下降。

结论。 我们的类别无关目标——针对SSR在潜在空间中向全局 最近代码进行蒸馏——明确地将监督与可能不正确的标签解耦，并将偏移区域与最兼容的语义混淆先验对齐，持续取得最佳性能。

W 跨数据集的类别映射

我们遵循SemanticKITTI（源 [A]）、Syn-激光雷达（源 [B]）和SemanticSTF（目标 [C]）所共享的标准19类评估协议。所有数据集均映射到ID为0–18的统一的19类分类体系；ID 255表示“忽略”，并被排除在损失和指标之外。

W.1 统一的19类分类体系

ID →类别：
0 汽车，1 自行车，2 摩托车，3 卡车，4 其他车辆，5 行人，6 骑自行车者，7 骑摩托车者，8 道路，9 停车区，

10 人行道，11 其他地面，12 建筑物，13 围栏，14 植被，15 树干，16 地形，17 杆，18 交通标志

W.2 [A] SEMANTICKITTI：原生标签与保留集合

标签到名称（原生）。

```
标签_名称_映射 = { 0: '未标注',
1: '离群点', 10: '汽车', 11: '自行车',
13: '公交车', 15: '摩托车', 16: '轨道
车辆', 18: '卡车', 20: '其他车辆',
30: '行人', 31: '骑自行车者', 32: '骑
摩托车者', 40: '道路', 44: '停车区',
48: '人行道', 49: '其他地面', 50: '建
筑物', 51: '围栏', 52: '其他结构',
60: '车道标线', 70: '植被', 71: '树干',
72: ' 地形 ', 80: ' 杆 ', 81: ' 交通标志 ',
99: ' 其他物体 ', 252: ' 移动汽车 ',
253: '移动的骑自行车者', 254: '移动
行人', 255: '移动的骑摩托车者',
256: '移动轨道车辆', 257: '移动公共
汽车', 258: '移动卡车', 259: '移动的
其他车辆' }
```

19 类协议保留的标签。

```
保留_标签 = [
'道路','人行道','停车区','其他地面','建筑物','汽车','卡车', '自行车','摩托车','其他车辆','植被','树
干','地形', '行人','骑自行车者','骑摩托车者','围栏','杆','交通标志']
```

注释。 移动类别（例如 252–259）会先映射到其静态对应类别（例如 moving-car→ 汽车），然后再应用 19 类筛选器，此过程遵循官方协议。不属于 19 类集合的类别将被指定为 255（忽略）。

W.3 [B] SYNLIDAR →统一的 19 类 ID

```
学习_映射: 0: 255 # 未标注 → 忽
略 1: 0 # 汽车 → 汽车 2: 3 # 皮卡
→ 卡车 3: 3 # 卡车 → 卡车
```

4: 4 # 公交车 → 其他车辆 5: 1 # 自行车 → 自行车 6: 2 # 摩托车 → 摩托车 7: 4 # 其他车辆 → 其他车辆 8: 8 # 道路 → 道路 9: 10 # 人行道 → 人行道 10: 9 # 停车区 → 停车区 11: 11 # 其他地面 → 其他地面 12: 5 # 女性 → 行人 13: 5 # 男性 → 行人 14: 5 # 儿童 → 行人 15: 5 # 群体 → 行人 16: 6 # 骑自行车者 → 骑自行车者 17: 7 # 骑摩托车者 → 骑摩托车者 18: 12 # 建筑物 → 建筑物 19: 255 # 其他结构 → 忽略 20: 14 # 植被 → 植被 21: 15 # 树干 → 树干 22: 16 # 地形 → 地形 23: 18 # 交通标志 → 交通标志 24: 17 # 杆 → 杆 25: 255 # 交通锥 → 忽略（不属于 19 类）26: 13 # 围栏 → 围栏 27: 255 # 垃圾桶 → 忽略 28: 255 # 电箱 → 忽略 29: 255 # 桌子 → 忽略 30: 255 # 椅子 → 忽略 31: 255 # 长椅 → 忽略 32: 255 # 其他物体 → 忽略

W.4 [C] SEMANTICSTF →统一的 19 类 ID

学习_映射 = { 0: 255, # 未标注 → 忽略 1: 0, # 汽车 2: 1, # 自行车 3: 2, # 摩托车 4: 3, # 卡车 5: 4, # 其他车辆 6: 5, # 行人 7: 6, # 骑自行车者 8: 7, # 骑摩托车者 9: 8, # 道路 10: 9, # 停车区 11: 10, # 人行道 12: 11, # 其他地面 13: 12, # 建筑物 14: 13, # 围栏 15: 14, # 植被 16: 15, # 树干 17: 16, # 地形 18: 17, # 杆 19: 18, # 交通标志 20: 255 # 无效 → 忽略 }

X 大语言模型与社会影响

大语言模型（LLM）的使用。在准备本稿件时，我们仅将大语言模型用于语言润色和写作辅助（例如，提升清晰度、语法和文风）。我们未使用大语言模型生成研究构想、实验结果、代码或分析，也未向大语言模型提供稿件文本以外的专有数据或敏感信息。所有技术内容、实验和结论均由作者完成并核实。

本文提出了一种在恶劣天气条件下实现域泛化激光雷达语义分割的方法。我们的贡献着重于通过增强感知潜在学习和语义偏移定位，提升模型对分布偏移的鲁棒性。目前，我们尚未发现所提方法本身会直接造成负面的社会影响。